VAN NOSTRAND REINHOLD COMPANY
Windsor House, 46 Victoria Street, London, S.W.1

INTERNATIONAL OFFICES
New York Cincinnati Toronto Melbourne

Library of Congress Catalog Card No. 78-141984

ISBN 0 442 06956 1

First published 1971

Printed in Great Britain by
Butler & Tanner Ltd
Frome and London

Contents

Preface

During the last ten or fifteen years, much progress has been made in the theory of non-self-adjoint compact linear operators on Hilbert space. The present volume is intended to provide a brief (and far from exhaustive) introduction to the subject. For the greater part, it is a revised version of a graduate course given at the University of Pennsylvania during the academic year 1964-65; some further material has been added subsequently, notably in sections 4.4 and 4.5.

A numeral in square brackets, in the text, refers to the bibliography included at the end of the book. Cross references to theorems, the principal definitions, etc., are given in decimal notation; § 3.2 is the second section of Chapter 3, and Lemma 3.2.6 is the sixth numbered item in that section. Equations are numbered consecutively, starting from (1), throughout each section; the twenty-third equation in § 3.2 is referred to as (23) within that section, but as 3.2(23) in subsequent sections.

I am indebted to the Trustees of the University of Pennsylvania, and to Professors Oscar Goldman and Richard Kadison, for arranging a visiting appointment in the Mathematics Department during the academic year 1964-65, when the lecture course which forms the basis of this volume was given. On two counts, my thanks are due to Dr Frank Smithies; first because, as my research supervisor during the middle 50s, he aroused my interest in many of the topics

in place of the more cumbersome

$$\{b \in B : \text{there exists } a \text{ in } A \text{ such that } P(a) \text{ and } f(a) = b\}.$$

Occasionally, it is convenient to refer to f as 'the mapping $a \to f(a)$'.

Suppose that K is either the real field or the complex field, and that X and Y are normed spaces over K; the norm in both spaces will be denoted by $\|\ \|$. We recall that a linear operator T from X into Y is continuous if and only if it is bounded, in the sense that the set $\{\|Tx\| : x \in X, \|x\| \leqslant 1\}$ of real numbers is bounded above. The set $\mathcal{B}(X, Y)$ of all bounded linear operators from X into Y is a vector space over K, and has a norm defined by

$$(1) \qquad \|T\| = \sup \{\|Tx\| : x \in X, \|x\| \leqslant 1\}$$

(equivalently, $\|T\|$ is the smallest real number C such that $\|Tx\| \leqslant C\|x\|$ for each x in X). If Y is a Banach space, then so is $\mathcal{B}(X, Y)$ [59: p. 163].

An element T of $\mathcal{B}(X, Y)$ is said to be *isometric* if $\|Tx\| = \|x\|$ for each x in X. If there is an isometric linear operator U from X onto Y, then X and Y are said to be *isometrically isomorphic*, and U is described as an *isometric isomorphism* from X onto Y.

A linear subspace M of a normed space X is itself a normed space. If M is a closed subspace, then the set X/M of all cosets $[x] = x+M$ of M in X is a normed space when the algebraic operations and norm are defined by

$$\alpha[x] = [\alpha x], \qquad [x]+[y] = [x+y],$$

$$\|[x]\| = \inf\{\|z\| : z \in [x]\}$$
$$= \inf\{\|x-y\| : y \in M\}.$$

We call X/M the *quotient space* of X modulo M. The mapping $x \to [x]$ is called the *canonical mapping* from X onto X/M, and is a norm decreasing (hence bounded) linear operator. If X is a Banach space, then so is X/M [59: pp. 104, 105].

If M and N are subspaces of a vector space X, we denote by $M+N$ the subspace $\{x+y : x \in M, y \in N\}$ of X.

THEOREM 1.1.1. *Suppose that M is a closed subspace of a normed space X and N is a finite-dimensional subspace of X. Then $M+N$ is a closed subspace of X.*

Proof. The canonical mapping Φ from X onto the quotient space X/M is continuous. Since $M+N$ is the inverse image under Φ of the finite-dimensional subspace $\Phi(N)$ of X/M, it suffices to show that $\Phi(N)$ is closed in X/M. For this, we refer to [59: p. 96].

If X is a normed space over K, then a linear operator from X into K is called a *linear functional* on X. Since K is a Banach space, it follows that the set $\mathcal{B}(X, K)$ of all bounded linear functionals on X is a Banach space (whether or not X is complete), with norm defined by

$$\|f\| = \sup\{|f(x)| : x \in X, \|x\| \leqslant 1\}.$$

We call $\mathcal{B}(X, K)$ the *dual space* of X, and denote it by X^*.

We assume that the reader is familiar with the Hahn-Banach theorem, and its immediate consequences, concerning the existence and properties of bounded linear functionals (see, for example, [59: pp. 185-188]). We recall one result of this type. Suppose that X is a normed space, $A \subseteq X$ and $B \subseteq X^*$. We can define closed subspaces A^0 of X^* and B_0 of X by the equations

(2) $$A^0 = \{f \in X^* : f(y) = 0 \ (y \in A)\},$$

(3) $$B_0 = \{x \in X : f(x) = 0 \ (f \in B)\}.$$

It is clear that $A \subseteq (A^0)_0$; and it follows easily, from the Hahn-Banach theorem, that $M = (M^0)_0$ when M is a closed subspace of X [59: Theorem 4.3-D, p. 186].

Given an element x of a normed space X, we can define a bounded linear functional $\hat{x}$ on the dual space X^* by the equation

$$\hat{x}(f) = f(x) \quad (f \in X^*).$$

The mapping $x \to \hat{x}$ is an isometric isomorphism from X onto a subspace X_0 of the second dual space $X^{**} (= (X^*)^*)$. We call X_0 *the canonical image of* X *in* X^{**}, and say that X is *reflexive* if $X_0 = X^{**}$ [59: p. 191].

If X and Y are normed spaces over K and B is a subset of $\mathcal{B}(X, Y)$ such that

$$\sup\{||T|| : T \in B\} < \infty,$$

then it follows trivially, from the inequality $|f(Tx)| \leqslant ||f|| \, ||T|| \, ||x||$, that

$$\sup\{|f(Tx)| : T \in B\} < \infty$$

whenever $f \in Y^*$ and $x \in X$. For Banach spaces, there is a converse result, the *principle of uniform boundedness.*

THEOREM 1.1.2. *Suppose that* X *and* Y *are Banach spaces over* K *and* $B \subseteq \mathcal{B}(X, Y)$. *If, for each* f *in* Y^* *and* x *in* X,

$$\sup\{|f(Tx)| : T \in B\} < \infty,$$

then

$$\sup\{||T|| : T \in B\} < \infty.$$

Proof. We content ourselves with showing that this theorem follows at once from very similar results proved, for example, in [59: pp. 201-204].

For a fixed element x of X,

$$\sup\{|f(Tx)| : T \in B\} < \infty$$

for each f in Y^*. By applying [59: Theorem 4.4-A, p. 202] to the subset $F = \{Tx : T \in B\}$ of Y, it follows that

$$\sup\{||Tx|| : T \in B\} < \infty.$$

This holds for each x in X so, by [59: Theorem 4.4-E, p. 204],

$$\sup\{||T|| : T \in B\} < \infty.$$

1.2. *Unordered summation*

Let f be a mapping from a set A into a (real or complex) Banach space X. We say that f *vanishes at infinity,* and write $f(a) \to 0$ as $a \to \infty$, if the following condition is satisfied: given any positive ϵ, the set $\{a \in A : ||f(a)|| \geqslant \epsilon\}$ is finite. When this is so, the set $\{a \in A : f(a) \neq 0\}$ is finite or countable, since it is the union of the finite sets $\{a \in A : ||f(A)|| \geqslant 1/n\}$ $(n = 1, 2, \ldots)$.

We introduce the idea of unordered summation in terms of the convergence of nets, following the terminology of [33: p. 65 et seq.]. The class $\mathcal{F}$ of all finite subsets of A is directed by the inclusion relation $\supseteq$, and we can define a mapping s from $\mathcal{F}$ into X by the equation

(1) $$s(F) = \sum_{a \in F} f(a) \quad (F \in \mathcal{F}).$$

We say that f is *summable* if the net $(s, \supseteq)$ converges to some element x of X; when this is so, we write

$$\sum_{a \in A} f(a) = x.$$

Thus f is summable if and only if there is an element x of X with the following property: given any positive ϵ, there exists $F(\epsilon)$ in $\mathfrak{F}$ such that

(2) $$\|x - \sum_{a \in F} f(a)\| < \epsilon \quad \text{whenever } F \in \mathfrak{F} \text{ and } F \supseteq F(\epsilon).$$

We sometimes express the fact that f is summable by saying that *the sum* $\sum_{a \in A} f(a)$ *exists*. Of course, if the set $B = \{a \in A : f(a) \neq 0\}$ is finite, then f is summable and x is the usual algebraic sum $\sum_{a \in B} f(a)$; for (2) is then satisfied if we take $F(\epsilon)$ to be B, for each positive ϵ.

Many of the expected manipulative properties of unordered sums are readily verified, and will be used in future without comment; for example, the reader should have no difficulty in supplying the missing proof of the following lemma.

LEMMA 1.2.1. *Suppose that f and g are mappings from a set A into a Banach space X, α and β are scalars, and T is a bounded linear operator from X into another Banach space Y.*

(i) If f and g are summable, then so is $\alpha f + \beta g$, and

$$\sum_{a \in A} \{\alpha f(a) + \beta g(a)\} = \alpha \sum_{a \in A} f(a) + \beta \sum_{a \in A} g(a).$$

(ii) The function $T \circ f$, defined by $(T \circ f)(a) = T(f(a))$, is summable, and

$$\sum_{a \in A} T(f(a)) = T(\sum_{a \in A} f(a)).$$

LEMMA 1.2.2. *Let f be a mapping from a set A into a Banach space X, and let $\mathfrak{F}$ denote the class of all finite subsets of A. Then f is summable if and only if it satisfies the following Cauchy condition: given any positive ϵ, there exists $G(\epsilon)$ in $\mathfrak{F}$ such that*

(3) $$\|\sum_{a \in F} f(a)\| < \epsilon \quad \textit{whenever } F \in \mathfrak{F} \textit{ and } F \subseteq A \sim G(\epsilon).$$

Proof. Suppose that f is summable and, for each positive ϵ, choose $F(\epsilon)$ in $\mathfrak{F}$ so that (2) is satisfied. If $F \in \mathfrak{F}$ and $F \subseteq A \sim F(\tfrac{1}{2}\epsilon)$, then both the sums $\sum_{a \in F(\frac{1}{2}\epsilon)} f(a)$ and $\sum_{a \in F \cup F(\frac{1}{2}\epsilon)} f(a)$ approximate x within $\tfrac{1}{2}\epsilon$; hence their difference, $\sum_{a \in F} f(a)$, has norm less than ϵ. This shows that the Cauchy condition (3) is satisfied, with $G(\epsilon) = F(\tfrac{1}{2}\epsilon)$.

Conversely, suppose that f has the Cauchy property stated in the lemma. If $F_j \in \mathfrak{F}$ and $F_j \supseteq G(\tfrac{1}{2}\epsilon)$ $(j = 1, 2)$ then (with the notation of (1))

$$\|s(F_j) - s(G(\tfrac{1}{2}\epsilon))\| = \|\sum_{a \in F_j \sim G(\frac{1}{2}\epsilon)} f(a)\| < \tfrac{1}{2}\epsilon,$$

and so $\|s(F_1) - s(F_2)\| < \epsilon$. This shows that $(s, \supseteq)$ is a Cauchy net; since the space X is complete, $(s, \supseteq)$ converges, and f is summable.

COROLLARY 1.2.3. *Let f be a mapping from a set A into a Banach space X. If f is summable, then $f(a) \to 0$ as $a \to \infty$.*

Proof. By taking for F a set with just one member, it follows from (3) that $\{a \in A : \|f(a)\| \geqslant \epsilon\}$ is a subset of $G(\epsilon)$ and is therefore finite, for each positive ϵ.

The preceding discussion applies, in particular, when X is the complex field. We now state some results about summability of

numerically-valued functions. The proofs are straightforward, and are left to the reader.

LEMMA 1.2.4. *Suppose that f is a complex-valued function defined on a set A, $\mathfrak{F}$ is the class of all finite subsets of A, and $s(F) = \sum_{a \in F} f(a)$ whenever $F \in \mathfrak{F}$.*

(i) $\sum_{a \in A} f(a)$ exists if and only if $\sum_{a \in A} |f(a)|$ exists; and when this is so, $|\sum_{a \in A} f(a)| \leqslant \sum_{a \in A} |f(a)|$.

(ii) If $f(a) > 0$ for each a in A, then $\sum_{a \in A} f(a)$ exists if and only if the set $\{s(F) : F \in \mathfrak{F}\}$ is bounded above; and when this is so,

$$\sum_{a \in A} f(a) = \sup \{s(F) : F \in \mathfrak{F}\}.$$

In view of part (ii) of Lemma 1.2.4, we sometimes express the fact that a non-negative real-valued function f is summable by writing $\sum_{a \in A} f(a) < \infty$.

LEMMA 1.2.5. *Suppose that f is a mapping from a set A into a Banach space X, and that $\sum_{a \in A} ||f(a)|| < \infty$. Then f is summable and*

$$||\sum_{a \in A} f(a)|| \leqslant \sum_{a \in A} ||f(a)||.$$

Proof. Since $\sum_{a \in A} ||f(a)|| < \infty$, it follows from Lemma 1.2.2 that there exists $G(\epsilon)$ in $\mathfrak{F}$ such that

$$\sum_{a \in F} ||f(a)|| < \epsilon \quad \text{whenever } F \in \mathfrak{F} \text{ and } F \subseteq A \sim G(\epsilon).$$

Hence

$$||\sum_{a \in F} f(a)|| < \epsilon \quad \text{wherever } F \in \mathfrak{F} \text{ and } F \subseteq A \sim G(\epsilon);$$

and, again by Lemma 1.2.2, f is summable. With $s(F)$ defined by (1), we have

$$||s(F)|| \leqslant \sum_{a \in F} ||f(a)|| \leqslant \sum_{a \in A} ||f(a)||.$$

Since the net $(s, \supseteq)$ converges to $\sum_{a \in A} f(a)$, it follows by taking limits that

$$||\sum_{a \in A} f(a)|| \leqslant \sum_{a \in A} ||f(a)||.$$

If A and B are sets, we denote by $A \times B$ the *Cartesian product* of A and B; that is, the set of all ordered pairs (a, b) with a in A and b in B.

LEMMA 1.2.6. *Suppose that f is a mapping from the Cartesian product $A \times B$ of two sets A and B into a Banach space X. If f is summable, then the repeated sum $\sum_{a \in A} [\sum_{b \in B} f(a, b)]$ exists, and is equal to $\sum_{(a, b) \in A \times B} f(a, b)$. If f is a non-negative real-valued function and the repeated sum exists, then f is summable.*

Note. By saying that the repeated sum exists we mean that, for each a in A, the sum $\sum_{b \in B} f(a, b) = h(a)$ exists, and the mapping h from A into X is summable. The proof of Lemma 1.2.6 is straightforward, and is left to the reader.

We conclude this section with a lemma which relates the concept of unordered summation to the usual notion of convergence of infinite series. Once again, the proof is left as an exercise for the reader.

LEMMA 1.2.7. *Let f be a mapping from a set A into a Banach space X, and let $B = \{a \in A : f(a) \neq 0\}$. Then the following two*

statements are equivalent:

(i) *f is summable.*

(ii) *B is finite or countable; if it is countable then, given any enumeration $\{b_j : j = 1, 2, \ldots\}$ of B, the infinite series $\Sigma\, f(b_j)$ converges.*

When these conditions are satisfied, the series in (ii) converge to $\sum_{a \in A} f(a)$.

1.3. *The spaces $\ell_p(A)$*

Given a set A and a positive real number p, we denote by $\ell_p(A)$ the set of all complex-valued functions f on A for which the sum $\sum_{a \in A} |f(a)|^p$ exists (in the sense of §1.2). If $f, g \in \ell_p(A)$, then $f+g \in \ell_p(A)$ and

$$\sum_{a \in A} |f(a)+g(a)|^p \leqslant 2^p\Big[\sum_{a \in A} |f(a)|^p + \sum_{a \in A} |g(a)|^p\Big].$$

To prove this it is sufficient, by Lemma 1.2.4 (ii), to establish the last inequality in the case where A is replaced by a finite subset F; and the required result then follows from the elementary inequalities

$$\begin{aligned}|f(a)+g(a)|^p &\leqslant [\max(|2f(a)|, |2g(a)|)]^p\\ &\leqslant |2f(a)|^p + |2g(a)|^p.\end{aligned}$$

Hence $\ell_p(A)$ is a complex vector space, since it is apparent that $\alpha f \in \ell_p(A)$ whenever $f \in \ell_p(A)$ and α is a complex number.

LEMMA 1.3.1. *(Minkowski's inequality). If $1 \leqslant p < \infty$ and $f, g \in \ell_p(A)$, then*

$$\Big[\sum_{a \in A} |f(a)+g(a)|^p\Big]^{1/p} \leqslant \Big[\sum_{a \in A} |f(a)|^p\Big]^{1/p} + \Big[\sum_{a \in A} |g(a)|^p\Big]^{1/p}.$$

We omit the proof of this result, remarking only that, once it has been established for summation over a *finite* set A, the extension for general A is an immediate consequence of Lemma 1.2.4 (ii).

It follows from Minkowski's inequality that the equation

$$||f||_p = \Big[\sum_{a \in A} |f(a)|^p\Big]^{1/p} \quad (f \in \ell_p(A))$$

defines a norm on $\ell_p(A)$. It is not difficult to verify that, with this norm, $\ell_p(A)$ is a Banach space when $1 \leqslant p < \infty$ [59: p. 100].

We denote by $\ell_\infty(A)$ the class of all bounded complex-valued functions on A. This is a complex vector space, and is a Banach space with respect to the norm $|| \;\; ||_\infty$ defined by

$$||f||_\infty = \sup\{|f(a)| : a \in A\} \quad (f \in \ell_\infty(A)).$$

Clearly $\ell_p(A) \subseteq \ell_q(A)$ $(0 < p \leqslant q \leqslant \infty)$.

If $p \geqslant 1$, $q \geqslant 1$ and $p^{-1} + q^{-1} = 1$, we say that p and q are *conjugate indices*, and describe each as the *conjugate index* of the other; for example, 1 and ∞ are conjugate indices, while 2 is its own conjugate index.

LEMMA 1.3.2. *(Hölder's inequality). If $1 \leqslant p \leqslant \infty$, q is the conjugate index of p, $f \in \ell_p(A)$ and $g \in \ell_q(A)$ then fg (the pointwise product, defined by $(fg)(a) = f(a)g(a)$) is in $\ell_1(A)$ and*

$$||fg||_1 \leqslant ||f||_p ||g||_q.$$

Once again we omit the proof, with the comment that the result is an immediate consequence of Lemma 1.2.4(ii), as soon as the inequality has been established in the case where A is a finite set.

Suppose that p and q are conjugate indices and $g \in \ell_q(A)$. For each f in $\ell_p(A)$, let

(1) $$\varphi_g(f) = \sum_{a \in A} f(a)g(a);$$

this sum exists, and furthermore $|\varphi_g(f)| \leqslant ||f||_p ||g||_q$, since $\sum_{a \in A} |f(a)g(a)| \leqslant ||f||_p ||g||_q$ by Lemma 1.3.2. It follows that φ_g is a bounded linear functional on the Banach space $\ell_p(A)$, and that $||\varphi_g|| \leqslant ||g||_q$. In fact, $||\varphi_g|| = ||g||_q$, and we have the following result [59: p. 194].

THEOREM 1.3.3. *Suppose that $1 \leqslant p < \infty$ and q is the conjugate index of p. When $g \in \ell_q(A)$, let φ_g be the element of $(\ell_p(A))^*$ defined by (1). Then the mapping $g \to \varphi_g$ is an isometric isomorphism from $\ell_q(A)$ onto $(\ell_p(A))^*$.*

1.4. *Analytic operator-valued functions*

Throughout § 1.4, X and Y are complex Banach spaces and $\mathcal{B}(X, Y)$ is the Banach space of all bounded linear operators from X into Y.

Suppose that D is an open subset of the complex field and φ is a function, defined on D, with values in $\mathcal{B}(X, Y)$. Given f in the dual space Y^* and x in X, the equation

(1) $$\varphi_{f,x}(\lambda) = f(\varphi(\lambda)x)$$

defines a complex-valued function $\varphi_{f,x}$ on D. We say that φ is *analytic on D* if, for every f in Y^* and x in X, $\varphi_{f,x}$ is analytic on D. If φ is defined and analytic on the whole complex field, it is said to be an *entire function*.

Many results concerning an operator-valued analytic function φ can be deduced by applying to each $\varphi_{f,x}$ the corresponding theorems

in classical complex variable theory. The proof of the following theorem illustrates this method.

THEOREM 1.4.1. *Suppose that $T_n \in \mathcal{B}(X, Y)$ $(n = 0, 1, 2, \ldots)$, r is a positive real number and D_r denotes the disc $\{\lambda : |\lambda| < r\}$. Then the following three conditions are equivalent.*

(i) For each f in Y^*, x in X and λ in D_r, the series $\sum_0^\infty \lambda^n f(T_n x)$ converges.

(ii) $$\limsup_{n \to \infty} ||T_n||^{1/n} \leqslant \frac{1}{r}.$$

(iii) For each λ in D_r, the series $\sum_0^\infty \lambda^n T_n$ converges with respect to the norm of $\mathcal{B}(X, Y)$.

When these conditions are satisfied, the function φ defined by

(2) $$\varphi(\lambda) = \sum_{n=0}^{\infty} \lambda^n T_n$$

is analytic on D_r, and its expansion in the form (2) is unique.

Proof. Suppose that (i) is satisfied. Given any real number a such that $0 < a < r$, the series $\Sigma\, a^n f(T_n x)$ converges and hence $a^n f(T_n x) \to 0$ as $n \to \infty$, whenever $f \in Y^*$ and $x \in X$. Thus

$$\sup \{|f(a^n T_n x)| : n = 0, 1, 2, \ldots\} < \infty$$

and, by the principle of uniform boundedness,

$$M = \sup \{||a^n T_n|| : n = 0, 1, 2, \ldots\} < \infty .$$

Hence

$$||T_n|| \leqslant M a^{-n}, \qquad ||T_n||^{1/n} \leqslant M^{1/n} a^{-1},$$

and

$$\limsup_{n \to \infty} ||T_n||^{1/n} \leqslant \frac{1}{a}$$

This has been proved whenever $0 < a < r$, and is therefore true also when $a = r$. Hence, (i) implies (ii).

If (ii) is satisfied then, by the elementary theory of numerical power series, $\Sigma\, |\lambda|^n\, ||T_n||$ converges whenever $|\lambda| < r$, and so $\Sigma\, \lambda^n T_n$ converges with respect to the norm on $\mathcal{B}(X, Y)$. Hence (ii) implies (iii).

Suppose that (iii) is satisfied, and φ is the function defined on D_r by (2). With $\varphi_{f,x}$ as in (1), for each f in Y^* and x in X,

$$\varphi_{f,x}(\lambda) = f(\varphi(\lambda)x) = \sum_{n=0}^{\infty} \lambda^n f(T_n x) \quad (\lambda \in D_r), \tag{3}$$

by (2) and the continuity of the linear functional $T \to f(Tx)$ on $\mathcal{B}(X, Y)$. This shows that (i) is satisfied, and that all the functions $\varphi_{f,x}$ are analytic on D_r. Thus φ is analytic on D_r.

If φ has another expansion, $\sum_0^\infty \lambda^n S_n$, of the form (2), then $\varphi_{f,x}$ has a second expansion $\sum_0^\infty \lambda^n f(S_n x)$ of the form (3). By the uniqueness of the classical Taylor expansion, $f(T_n x) = f(S_n x)$, for all f in Y^* and x in X. Thus $T_n = S_n$ $(n = 0, 1, 2, \ldots)$, and the final clause of the theorem is proved.

If φ is analytic on an open subset D of the complex plane, and takes values in $\mathcal{B}(X, Y)$, then φ is continuous on D when $\mathcal{B}(X, Y)$ has the norm topology. To prove this, we have to show that, if $\lambda_0 \in D$, then $||\varphi(\lambda) - \varphi(\lambda_0)|| \to 0$ as $\lambda \to \lambda_0$. Now D contains a compact disc D_0 with centre at λ_0. For each f in Y^* and x in X, the function $\varphi_{f,x}$ in (1) is analytic on D; its derivative $\varphi'_{f,x}$ is analytic on D and thus bounded on D_0. If $\lambda \in D_0$, $\varphi_{f,x}(\lambda) - \varphi_{f,x}(\lambda_0)$ is the integral of $\varphi'_{f,x}$ along the line segment from λ_0 to λ; so

$$|\varphi_{f,x}(\lambda) - \varphi_{f,x}(\lambda_0)| \leqslant M_{f,x}\, |\lambda - \lambda_0|,$$

where $M_{f,x}$ is an upper bound of $|\varphi'_{f,x}|$ on D_0. The last inequality can be rewritten as

$$\left| f\left(\frac{1}{\lambda-\lambda_0}\, [\varphi(\lambda)-\varphi(\lambda_0)]x\right) \right| \leqslant M_{f,x} \quad (\lambda \in D_0,\ \lambda \neq \lambda_0).$$

By the principle of uniform boundedness, there is a constant M such that

$$\left|\left| \frac{1}{\lambda-\lambda_0}\, [\varphi(\lambda)-\varphi(\lambda_0)] \right|\right| \leqslant M \quad (\lambda \in D_0,\ \lambda \neq \lambda_0).$$

Thus $||\varphi(\lambda)-\varphi(\lambda_0)|| \leqslant M\, |\lambda-\lambda_0| \to 0$, as $\lambda \to \lambda_0$.

Suppose that φ is an entire function, with values in $\mathcal{B}(X, Y)$. If P is a function which is defined and positive for all complex λ, we write $\varphi(\lambda) = O(P(\lambda))$ if there is a constant M such that $||\varphi(\lambda)|| \leqslant MP(\lambda)$ for all λ. With this convention, φ is said to be of *finite order* if there is a positive real number σ such that

$$\varphi(\lambda) = O(\exp(|\lambda|^\sigma)); \tag{4}$$

when this is so, the infimum of the set of all such positive σ is called the *order* of φ. If φ has finite order ρ, then φ is said to be of *minimum type* if

$$\varphi(\lambda) = O(\exp(\epsilon|\lambda|^\rho)) \tag{5}$$

for every positive ϵ. Note that, if φ satisfies one or other of the conditions (4), (5), then the same is true of the complex-valued functions $\varphi_{f,x}$ $(f \in Y^*,\ x \in X)$ defined by (1).

We conclude this section with two results from classical complex variable theory which will be needed in later chapters.

Suppose that φ is a (complex-valued) entire function, with $\varphi(0) \neq 0$, which has finite order ρ. The zeros of φ, counted according to their multiplicities, can be arranged as a sequence $\lambda_1, \lambda_2, \lambda_3, \ldots,$ with

$$|\lambda_1| \leqslant |\lambda_2| \leqslant |\lambda_3| \leqslant \ldots$$

It can be shown that

$$\Sigma\, |\lambda_n|^{-\alpha} < \infty \tag{6}$$

whenever $\alpha > \rho$. If the smallest positive integer α satisfying (6) is p+1, then p is called the *genus of the zeros* of φ. The product

$$P(\lambda) = \Pi \left(1-\frac{\lambda}{\lambda_n}\right) \exp\left(\frac{\lambda}{\lambda_n} + \frac{\lambda^2}{2\lambda_n^2} + \ldots + \frac{\lambda^p}{p\lambda_n^p}\right) \tag{7}$$

(in which the exponential does not appear if $p = 0$) converges uniformly on each bounded subset of the complex field, and is called the *canonical product formed with the zeros of* φ (if φ has no zeros, $P(\lambda) = 1$). For these facts, and for a proof of the following result, we refer to [60: pp. 246–51].

THEOREM 1.4.2. *(Hadamard's factorization theorem). If φ is an entire function with finite order ρ, and $\varphi(0) \neq 0$, then*

$$\varphi(\lambda) = e^{Q(\lambda)} P(\lambda),$$

where $Q(\lambda)$ is a polynomial of degree not greater than ρ and P is the canonical product formed with the zeros of φ.

THEOREM 1.4.3. *Suppose that φ is an entire function, $\varphi(\lambda) \neq 0$ for all complex λ, $|\varphi(\lambda)| = 1$ when $\operatorname{Im} \lambda = 0$, $|\varphi(\lambda)| > 1$ when $\operatorname{Im} \lambda > 0$, and $\varphi(0) = 1$. Then there is a positive real number a such that $\varphi(\lambda) = \exp(-ia\lambda)$.*

Proof. Since $\varphi'(\lambda)/\varphi(\lambda)$ is an entire function, the integral

$$g(\lambda) = i\int_0^\lambda \frac{\varphi'(z)}{\varphi(z)}\, dz$$

is the same for all rectifiable arcs from 0 to λ, and g is an entire function satisfying $g(0) = 0$, $g'(\lambda) = i\varphi'(\lambda)/\varphi(\lambda)$. The function $\varphi(\lambda)\exp(ig(\lambda))$ has zero derivative, and is therefore identically equal to $\varphi(0)\exp(ig(0))$ (= 1). Thus

$$\varphi(\lambda) = \exp(-ig(\lambda)).$$

The conditions on $|\varphi(\lambda)|$ imply that $g(\lambda)$ is real when λ is real, and that $\operatorname{Im} g(\lambda) > 0$ when $\operatorname{Im} \lambda > 0$. If we define

$$f_r(\theta) = \operatorname{Im} g(re^{i\theta}) \tag{8}$$

for real θ and positive r, then

$$f_r(\theta) > 0 \quad (0 < \theta < \pi,\ r > 0). \tag{9}$$

The Taylor expansion

$$g(\lambda) = a_1\lambda + a_2\lambda^2 + \ldots \tag{10}$$

of g is valid for all complex λ. Since $g(\lambda)$ is real when λ is real, it follows that a_n $(=g^{(n)}(0)/n!)$ is real. By (8) and (10),

$$f_r(\theta) = a_1 r \sin\theta + a_2 r^2 \sin 2\theta + \ldots$$

For a fixed positive r, this series converges uniformly for θ in the compact interval $[0,\pi]$, so

$$a_n r^n = \frac{2}{\pi}\int_0^\pi f_r(\theta)\sin n\theta\, d\theta. \tag{11}$$

It is easily proved, by induction on n, that

$$|\sin n\theta| \leqslant n \sin \theta \quad (0 < \theta < \pi;\ n = 1, 2, \ldots).$$

This, with (9) and (11), gives

$$|a_n r^n| \leqslant \frac{2}{\pi}\int_0^\pi f_r(\theta)\, |\sin n\theta|\, d\theta$$

$$\leqslant \frac{2n}{\pi}\int_0^\pi f_r(\theta) \sin\theta\, d\theta = na_1 r.$$

Thus $|a_n| \leqslant na_1 r^{1-n}$; by letting $r \to \infty$, we have $a_n = 0$ $(n > 1)$. It follows from (10) that $g(\lambda) = a_1\lambda$, and so $\varphi(\lambda) = \exp(-ia_1\lambda)$. Since $|\varphi(\lambda)| > 1$ when $\operatorname{Im}\lambda > 0$, we have $a_1 > 0$.

1.5. *Elementary spectral theory*

Throughout § 1.5, X is a complex Banach space and I denotes the identity operator on X. We write $\mathcal{B}(X)$, instead of $\mathcal{B}(X, X)$, for the Banach space of all bounded linear operators from X into itself. If $S, T \in \mathcal{B}(X)$, then $ST \in \mathcal{B}(X)$ and $||ST|| \leqslant ||S||\,||T||$. Thus $\mathcal{B}(X)$ is an associative linear algebra; furthermore, for a fixed S in $\mathcal{B}(X)$, the mappings $T \to ST$ and $T \to TS$ are bounded linear operators on the Banach space $\mathcal{B}(X)$.

The *spectral radius* $r(T)$ of an element T of $\mathcal{B}(X)$ is defined by $r(T) = \inf\{||T^n||^{1/n}\colon n = 1, 2, \ldots\}$. It is clear that

(1) $$0 \leqslant r(T) \leqslant ||T||;$$

we assert also that

(2) $$r(T) = \lim_{n\to\infty} ||T^n||^{1/n}.$$

To prove (2), choose any positive integer k and let $\alpha = ||T^k||^{1/k}$. Each $n = 1, 2, \ldots$ can be expressed as $n = q(n)k + r(n)$, where

$q(n) \geqslant 0$ and $0 \leqslant r(n) < k$. Then

$$\begin{aligned} ||T^n|| &= ||(T^k)^{q(n)} T^{r(n)}|| \\ &\leqslant ||T^k||^{q(n)}\, ||T||^{r(n)} \\ &= \alpha^{k\,q(n)}\, ||T||^{r(n)} \\ &= \alpha^{n-r(n)}\, ||T||^{r(n)} \end{aligned}$$

Thus $||T^n||^{1/n} \leqslant \alpha^{1-r(n)/n}\, ||T||^{r(n)/n}$ and, since $r(n)$ is bounded,

$$\limsup_{n\to\infty} ||T^n||^{1/n} \leqslant \alpha = ||T^k||^{1/k}.$$

It follows that

$$\begin{aligned} \limsup_{n\to\infty} ||T^n||^{1/n} &\leqslant \inf\{||T^k||^{1/k} : k = 1, 2, \ldots\} \\ &= r(T) \leqslant \liminf_{n\to\infty} ||T^n||^{1/n}, \end{aligned}$$

and (2) is proved.

By Theorem 1.4.1 and equation (2), the series $\Sigma\, \lambda^n T^n$ converges, with respect to the norm on $\mathcal{B}(X)$, whenever $|\lambda|\, r(T) < 1$; this is so, in particular, if $|\lambda|\, ||T|| < 1$. Since multiplication by λT is a bounded linear operator on $\mathcal{B}(X)$, the operator

$$S = \sum_{n=0}^{\infty} \lambda^n T^n$$

satisfies

$$\lambda TS = \lambda ST = \sum_{n=0}^{\infty} \lambda^{n+1} T^{n+1} = S - I,$$

$$(I - \lambda T)S = S(I - \lambda T) = I.$$

Hence $I-\lambda T$ has an inverse in $\mathcal{B}(X)$ when $|\lambda|\ r(T) < 1$, and

$$(3)\qquad (I-\lambda T)^{-1} = \sum_{n=0}^{\infty} \lambda^n T^n \quad (|\lambda|\ r(T) < 1).$$

It is frequently convenient to make a change of variable, $\lambda = \mu^{-1}$, and to consider the existence of $(\mu I-T)^{-1}$. The *resolvent set* of T is the set $\rho(T)$ of all complex numbers μ for which $\mu I-T$ has an inverse in $\mathcal{B}(X)$; its complement is called the *spectrum* of T and is denoted by $\sigma(T)$. The set $\sigma_p(T)$, consisting of all complex numbers μ such that $\mu I-T$ is not one to one, is a subset of $\sigma(T)$ and is called the *point spectrum* of T; elements of $\sigma_p(T)$ are called *eigenvalues* of T. If $\mu \in \sigma_p(T)$, n is a positive integer, and x is a non-zero element of X which satisfies $(\mu I-T)^n x = 0$, then x is called a *principal vector* of T associated with the eigenvalue μ; in the case $n = 1$, x is described as an *eigenvector* of T. The operator-valued function $R(\mu, T)$, defined on $\rho(T)$ by

$$R(\mu, T) = (\mu I-T)^{-1},$$

is called the *resolvent* of T. It follows from (3) that

$$(4)\qquad \mu \in \rho(T), \qquad R(\mu, T) = \sum_{n=0}^{\infty} \mu^{-n-1} T^n \ (|\mu| > r(T));$$

and so

$$(5)\qquad \sigma(T) \subseteq \{\mu : |\mu| \leqslant r(T)\}.$$

THEOREM 1.5.1. *Suppose that X is a complex Banach space (not consisting of the zero vector only) and $T \in \mathcal{B}(X)$. Then $\sigma(T)$ is a compact non-empty subset of the complex plane, and*

$$(6)\qquad \max\{|\mu| : \mu \in \sigma(T)\} = r(T).$$

The resolvent set $\rho(T)$ is open, and $R(\mu, T)$ is analytic on $\rho(T)$.

Proof. If $\mu_0 \in \rho(T)$ and $R(\mu_0, T) = R_0$, then

$$\mu I-T = \mu_0 I-T+(\mu-\mu_0)I$$
$$= (\mu_0 I-T)(I+(\mu-\mu_0)R_0).$$

If $|\mu-\mu_0|\ \|R_0\| < 1$, then $I+(\mu-\mu_0)R_0$ has an inverse, $\Sigma\,(\mu_0-\mu)^n R_0^n$; thus $\mu I-T$ has an inverse $R(\mu, T)$ given by

$$R(\mu, T) = \sum_{n=0}^{\infty} (\mu_0-\mu)^n R_0^{n+1},$$

and so $\mu \in \rho(T)$. It follows that $\rho(T)$ contains the disc $\{\mu : |\mu-\mu_0|\ \|R_0\| < 1\}$, and that $R(\mu, T)$ is analytic on this disc. Hence $\rho(T)$ is open, and $R(\mu, T)$ is analytic on $\rho(T)$.

The complement $\sigma(T)$ of $\rho(T)$ is closed; it is bounded, by (5), and is therefore compact. Let r be any real number such that

$$(7)\qquad r > 0, \qquad \sigma(T) \subseteq \{\mu : |\mu| \leqslant r\}.$$

We shall prove that these conditions imply that $r \geqslant r(T)$. If λ is a complex number such that $0 < |\lambda| < r^{-1}$, then $|\lambda^{-1}| > r$ and, by (7), $\lambda^{-1} \in \rho(T)$. Hence $I-\lambda T$ has an inverse,

$$(I-\lambda T)^{-1} = [\lambda(\lambda^{-1}I-T)]^{-1} = \lambda^{-1}R(\lambda^{-1}, T).$$

This shows that $(I-\lambda T)^{-1}$ exists and is analytic throughout the punctured disc $\{\lambda : 0 < |\lambda| < r^{-1}\}$; by (3), it is also analytic on a neighbourhood of 0, and is therefore analytic on the whole of the disc $\{\lambda : |\lambda| < r^{-1}\}$. For each f in X^* and x in X, the complex-valued function $\varphi_{f,x}(\lambda) = f((I-\lambda T)^{-1}x)$ is analytic when $|\lambda| < r^{-1}$; so its Taylor expansion, which by (3) is $\sum_0^{\infty} \lambda^n f(T^n x)$, converges when $|\lambda| < r^{-1}$. From the equivalence of conditions (i)

and (ii) in Theorem 1.4.1, it follows that

$$r \geqslant \limsup_{n\to\infty} \|T^n\|^{1/n} = r(T).$$

This shows that, if r satisfies (7), then $r \geqslant r(T)$.

Suppose that $\sigma(T)$ is empty. Given any positive r, (7) is satisfied and therefore $r(T) \leqslant r$; thus $r(T) = 0$. However, since $0 \notin \sigma(T)$, T has an inverse S in $\mathcal{B}(X)$. Since X does not consist of the zero vector only,

$$1 = \|I\| = \|(ST)^n\| = \|S^n T^n\| \leqslant \|S\|^n \|T^n\|,$$

and so $\|T^n\|^{1/n} \geqslant \|S\|^{-1}$ $(n = 1, 2, \ldots)$. Thus $r(T) \geqslant \|S\|^{-1} > 0$, contradicting our previous conclusion that $r(T) = 0$. This shows that $\sigma(T)$ is not empty. By (5),

$$\max\{|\mu| : \mu \in \sigma(T)\} \leqslant r(T). \tag{8}$$

If strict inequality occurs, we can choose r so that (7) is satisfied and $r < r(T)$, contradicting our previous conclusion that $r \geqslant r(T)$ for each such r. Hence, equality occurs in (8).

An element T of $\mathcal{B}(X)$ is said to be *quasi-nilpotent* if $r(T) = 0$; by Theorem 1.5.1, this occurs if and only if $\sigma(T)$ consists of the single point 0.

If $T \in \mathcal{B}(X)$, and $p(\mu) = a_0 + a_1\mu + \ldots + a_n\mu^n$ is a polynomial with complex coefficients, then we denote by $p(T)$ the element $a_0 I + a_1 T + \ldots + a_n T^n$ of $\mathcal{B}(X)$.

THEOREM 1.5.2. *(Spectral mapping theorem). If X is a complex Banach space, $T \in \mathcal{B}(X)$ and p is a polynomial with complex coefficients, then*

$$\sigma(p(T)) = \{p(\mu) : \mu \in \sigma(T)\}.$$

Proof. We begin by remarking that, if $S_j \in \mathcal{B}(X)$ and $S_i S_j = S_j S_i$ $(i, j = 1, \ldots, n)$, then the product $S_1 S_2 \ldots S_n$ has an inverse if and only if each S_j has an inverse; the proof is straightforward, and is left to the reader. The result of the theorem is obvious if p is a constant, so we may assume that

$$p(\mu) = a_0 + a_1\mu + \ldots + a_n\mu^n,$$

where $n \geqslant 1$ and $a_n \neq 0$.

If $\mu_0 \in \sigma(T)$, then

$$p(\mu_0)I - p(T) = \sum_{j=0}^{n} a_j(\mu_0^j I - T^j)$$
$$= (\mu_0 I - T)q(T)$$

for some polynomial q. Since $\mu_0 I - T$ has no inverse, neither does $p(\mu_0)I - p(T)$, and therefore $p(\mu_0) \in \sigma(p(T))$.

Next, suppose that ν is a complex number, not of the form $p(\mu)$ with μ in $\sigma(T)$. Then $p(\mu) - \nu$ can be factorized as

$$p(\mu) - \nu = a_n(\mu - \mu_1)(\mu - \mu_2) \ldots (\mu - \mu_n),$$

where none of the complex numbers $\mu_1, \ldots, \mu_n$ is in $\sigma(T)$. It follows that

$$p(T) - \nu I = a_n(T - \mu_1 I)(T - \mu_2 I) \ldots (T - \mu_n I);$$

and since the (commuting) operators on the right-hand side all have inverses, so does $\nu I - p(T)$. Thus, $\nu \notin \sigma(p(T))$.

Suppose that $T \in \mathcal{B}(X)$. By a *closed invariant* subspace of T we mean a closed subspace M of X with the property that $Tx \in M$

whenever $x \in M$; we also describe this property by saying that M is *invariant under* T. There are at least two such subspaces, the trivial ones $\{0\}$ and X; any other ones are described as *proper* closed invariant subspaces. If X has finite dimension greater than 1, the existence of an eigenvector of T ensures that T has a proper closed invariant subspace. For infinite-dimensional Banach spaces (even Hilbert spaces) it is not known whether or not each T in $\mathcal{B}(X)$ has a proper closed invariant subspace, although affirmative results are known for certain classes of operators.

If $T \in \mathcal{B}(X)$, then $\rho(T)$ is open and therefore each of its connected components is open; since $\rho(T)$ has compact complement, exactly one of its connected components, which we denote by $\rho_\infty(T)$, is unbounded.

THEOREM 1.5.3. *Suppose that X is a complex Banach space, $T \in \mathcal{B}(X)$ and M is a closed invariant subspace of T. Then M is invariant under $(\mu I-T)^{-1}$ whenever $\mu \in \rho_\infty(T)$, the unbounded connected component of $\rho(T)$.*

Proof. Let M^0 be the closed subspace of X^* defined as in 1.1(2). Given f in M^0 and x in M, the function φ defined by

$$\varphi(\mu) = f((\mu I-T)^{-1}x)$$

is analytic on $\rho_\infty(T)$. Since $T^n x \in M$, it follows that $f(T^n x) = 0$ $(n = 0, 1, 2, \ldots)$. By (4), $\mu \in \rho_\infty(T)$ and

$$\varphi(\mu) = \sum_{n=0}^{\infty} \mu^{-n-1} f(T^n x) = 0$$

whenever $|\mu| > r(T)$. Since φ vanishes on a non-empty open subset of the connected set $\rho_\infty(T)$, it follows by analytic continuation that φ vanishes throughout $\rho_\infty(T)$.

We have shown that, if $x \in M$ and $\mu \in \rho_\infty(T)$, then

$$f((\mu I-T)^{-1}x) = 0 \quad (f \in M^0).$$

Thus $(\mu I-T)^{-1}x \in (M^0)_0 = M$.

If $T \in \mathcal{B}(X)$, M is a closed invariant subspace of T, and T^M denotes the restriction of T to M, then $T^M \in \mathcal{B}(M)$ and $\|T^M\| \leqslant \|T\|$. For x in X, let $[x]$ be the coset $x+M$ $(\in X/M)$. If $[x] = [y]$, then $x-y \in M$, $Tx-Ty \in M$, and thus $[Tx] = [Ty]$. It follows that the mapping $T_M : [y] \to [Ty]$ is well-defined; and it is readily verified that $T_M \in \mathcal{B}(X/M)$ and $\|T_M\| \leqslant \|T\|$.

1.6. *Hilbert spaces*

We assume that the reader is familiar with the elementary theory of inner product spaces (see, for example, [26: Chapter 3]).

Let $\mathcal{H}$ be a complex inner product space, and denote by $\langle x, y\rangle$ the inner product of vectors x and y in $\mathcal{H}$. The equation

$$\|x\| = \langle x, x\rangle^{1/2} \tag{1}$$

defines a norm on $\mathcal{H}$, and the *Cauchy-Schwarz inequality* asserts that

$$|\langle x, y\rangle| < \|x\|\|y\| \quad (x, y \in \mathcal{H}). \tag{2}$$

This implies that the inner product is (jointly) continuous in both variables, with respect to the norm topology on $\mathcal{H}$. In particular, for each fixed y in $\mathcal{H}$, the mapping $x \to \langle x, y\rangle$ is a bounded linear functional on $\mathcal{H}$.

A complex inner product space which is complete with respect to the norm in (1) is called a *Hilbert space*. We emphasize that, in this book, the term Hilbert space is used only in connection with *complex* vector spaces.

If $\mathcal{H}$ is the Banach space $\ell_2(A)$ described in § 1.3, the equation

$$\langle f, g\rangle = \sum_{a\in A} f(a)\overline{g(a)} \quad (f, g \in \ell_2(A))$$

defines an inner product on $\mathcal{H}$; the sum exists, by Lemma 1.3.2. The norm derived from this inner product, as in (1), is the usual one, $\|\ \|_2$, on $\ell_2(A)$. Thus $\ell_2(A)$ is a Hilbert space.

Suppose that $\mathcal{H}$ is a Hilbert space, A and B are subsets of $\mathcal{H}$, and $x, y \in \mathcal{H}$. We say that x is *orthogonal* to y if $\langle x, y\rangle = 0$. If $\langle x, b\rangle = 0$ whenever $b \in B$, we say that x is orthogonal to B; the set of all vectors which are orthogonal to B is denoted by $B^{\perp}$. If $\langle a, b\rangle = 0$ whenever $a \in A$ and $b \in B$, we say that A is orthogonal to B.

Since the mapping $x \to \langle x, b\rangle$ is a bounded linear functional on $\mathcal{H}$, for each b, it is easily verified that $B^{\perp}$ is a closed subspace of $\mathcal{H}$. Let M be the closed subspace generated by B. If $x \in B^{\perp}$, then $\langle x, b\rangle = 0$ for each b in B, and the conjugate linearity and continuity of the mapping $c \to \langle x, c\rangle$ implies that $\langle x, c\rangle = 0$ for each c in M. Thus, $B^{\perp} \subseteq M^{\perp}$; since the reverse inclusion is apparent, $B^{\perp} = M^{\perp}$. Note also that $B \cap B^{\perp} = \{0\}$ since, if $x \in B \cap B^{-}$, we have $0 = \langle x, x\rangle = \|x\|^2$.

Suppose that $\mathcal{H}$ is a Hilbert space and $A \subseteq \mathcal{H}$. We say that A is an *orthogonal set* if $\langle a, b\rangle = 0$ whenever $a, b \in A$ and $a \neq b$; if, in addition, $\|a\| = 1$ for each a in A, we say that A is an *orthonormal set*. An orthonormal set A is linearly independent (by which, we mean that every finite subset is linearly independent); for if $a_1, \ldots, a_n \in A$, $\alpha_1, \ldots, \alpha_n$ are complex numbers and $\sum \alpha_j a_j = 0$, then

$$0 = \langle \sum \alpha_j a_j, a_k\rangle = \alpha_k$$

for each $k = 1, \ldots, n$.

In the following lemmas, we make use of the concept of unordered summation, which was described in § 1.2.

LEMMA 1.6.1. *Suppose that A is an orthogonal set in a Hilbert space $\mathcal{H}$.*

(i) $\sum_{a\in A} a$ exists if an only if $\sum_{a\in A} \|a\|^2 < \infty$. When this is so, $\|\sum_{a\in A} a\|^2 = \sum_{a\in A} \|a\|^2$.

(ii) If A is orthonormal and f is a complex-valued function on A, then $\sum_{a\in A} f(a)a$ exists if and only if $\sum_{a\in A} |f(a)|^2 < \infty$. When this is so, $\|\sum_{a\in A} f(a)a\|^2 = \sum_{a\in A} |f(a)|^2$.

Proof. It suffices to prove (i), since (ii) then follows at once. For any finite subset F of A,

$$(3) \qquad \|\sum_{a\in F} a\|^2 = \langle \sum_{a\in F} a, \sum_{b\in F} b\rangle = \sum_{a, b\in F} \langle a, b\rangle = \sum_{a\in F} \|a\|^2.$$

It now follows, from the Cauchy criterion (Lemma 1.2.2) that $\sum_{a\in A} a$ exists if and only if $\sum_{a\in A} \|a\|^2 < \infty$. When this is so, the infinite sums in (i) are limits of the appropriate nets of finite subsums; since the norm is a continuous function on $\mathcal{H}$, it follows from (3) that

$$\|\sum_{a\in A} a\|^2 = \sum_{a\in A} \|a\|^2.$$

LEMMA 1.6.2. *If A is an orthonormal set in a Hilbert space $\mathcal{H}$, and $x \in \mathcal{H}$, then*

(i) $\sum_{a\in A} |\langle x, a\rangle|^2 \leqslant \|x\|^2$ (Bessel's inequality);

(ii) the sum $x_1 = \sum_{a \in A} \langle x, a\rangle a$ *exists, and* $x - x_1 \in A^{\perp}$;

(iii) $||x - x_1||^2 = ||x||^2 - \sum_{a \in A} |\langle x, a\rangle|^2$.

Proof. If F is a finite subset of A, and $g(a) = \langle x, a\rangle$ for each a in A, then

$$\begin{aligned} ||x - \sum_{a \in F} \langle x, a\rangle a||^2 &= \langle x - \sum_{a \in F} g(a)a,\ x - \sum_{b \in F} g(b)b\rangle \\ &= ||x||^2 - \sum_{a \in F} g(a)\langle a, x\rangle \\ &\quad - \sum_{b \in F} \overline{g(b)}\langle x, b\rangle \\ &\quad + \sum_{a,\, b \in F} g(a)\overline{g(b)}\langle a, b\rangle \\ &= ||x||^2 - \sum_{a \in F} |\langle x, a\rangle|^2. \end{aligned}$$

Thus, for each finite subset F of A,

$$\sum_{a \in F} |\langle x, a\rangle|^2 = ||x||^2 - ||x - \sum_{a \in F} \langle x, a\rangle a||^2 \leqslant ||x||^2. \tag{4}$$

This proves (i), and the existence of the sum x_1 in (ii) now follows from Lemma 1.6.1. Since the mapping $y \to \langle y, b\rangle$ is a bounded linear functional on $\mathcal{H}$, for each b in A, we have

$$\langle x_1, b\rangle = \sum_{a \in A} \langle x, a\rangle\langle a, b\rangle = \langle x, b\rangle \quad (b \in A),$$

so $x - x_1 \in A^{\perp}$. This proves (ii); and (iii) follows from (4) since the norm is a continuous function on $\mathcal{H}$.

THEOREM 1.6.3. *If A is an orthonormal set in a Hilbert space $\mathcal{H}$, the following five conditions are equivalent.*

(i) $A^{\perp} = \{0\}$.

(ii) For each x in $\mathcal{H}$, $x = \sum_{a \in A} \langle x, a\rangle a$.

(iii) For each x and y in $\mathcal{H}$, $\langle x, y\rangle = \sum_{a \in A} \langle x, a\rangle\langle a, y\rangle$.

(iv) For each x in $\mathcal{H}$, $||x||^2 = \sum_{a \in A} |\langle x, a\rangle|^2$.

(v) The closed subspace generated by A is the whole of $\mathcal{H}$.

Proof. Suppose that (i) is satisfied. If $x \in \mathcal{H}$ then, by Lemma 1.6.2(ii), the sum $\sum_{a \in A} \langle x, a\rangle a$ exists, and

$$x - \sum_{a \in A} \langle x, a\rangle a \in A^{\perp} = \{0\}.$$

Hence, (i) implies (ii). Since the mapping $z \to \langle z, y\rangle$ is a bounded linear functional on $\mathcal{H}$, it is apparent that (ii) implies (iii). By taking $y = x$ in (iii), it follows that (iii) implies (iv).

Suppose that (iv) is satisfied. Given x in $\mathcal{H}$, Lemma 1.6.2(iii) implies that $x = \sum_{a \in A} \langle x, a\rangle a$, so x lies in the closed subspace M generated by A. Thus $M = \mathcal{H}$; so (iv) implies (v). Finally, if (v) is satisfied, then $A^{\perp} = \mathcal{H}^{\perp} = \{0\}$; so (v) implies (i).

Let $\mathcal{H}$ be a Hilbert space and suppose (to avoid trivialities) that $\mathcal{H}$ does not consist of the zero vector only. Then there exist non-empty orthonormal sets A_0 in $\mathcal{H}$. A simple argument based on Zorn's Lemma shows that each such A_0 is contained in an orthonormal set A which is *maximal*, in the sense that A is not a proper subset of any orthonormal set B. It is easily verified that an orthonormal set A is maximal if and only if $A^{\perp} = \{0\}$. By Theorem 1.6.3, A is maximal if and only if it verifies any one (and hence all) of the conditions (i), . . ., (v) listed in that theorem. In view of (ii), a maximal orthonormal set in $\mathcal{H}$ is called an *orthonormal basis* of $\mathcal{H}$; the equality in (iii) is known as *Parseval's equation*.

In the preceding paragraph we have proved the following result.

THEOREM 1.6.4. *If $\mathcal{H}$ is a Hilbert space which does not consist of the zero vector only, then $\mathcal{H}$ has an orthonormal basis. Furthermore, every orthonormal set in $\mathcal{H}$ is contained in an orthonormal basis.*

The *dimension* of a Hilbert space $\mathcal{H}$, denoted by $\dim(\mathcal{H})$, is defined to be the cardinal number of an orthonormal basis A of $\mathcal{H}$; it can be shown [59: p. 117] that this does not depend on the choice of A. Hilbert spaces with the same dimension are isometrically isomorphic; separable infinite-dimensional Hilbert spaces have countable orthonormal bases and are isomorphic to $\ell_2(N)$, where N is the set of all positive integers [59: pp. 120, 121].

If $\mathcal{H}$ is a Hilbert space and J is any set with cardinal number equal to $\dim(\mathcal{H})$, then every orthonormal basis of $\mathcal{H}$ can be indexed by J, and so expressed in the form $\{\varphi_j : j \in J\}$. If A is an orthonormal set in $\mathcal{H}$ then, since A is contained in an orthonormal basis, it can be expressed as $\{\psi_k : k \in K\}$, where K is a subset of J. Of course, this can be done in many ways: when an orthonormal set is given in this form, we shall use the term 'orthonormal *system*', to indicate that a particular indexing is specified. We shall reserve the phrase 'orthonormal *sequence*' for the case in which K is $\{1, 2, 3, \ldots\}$, or $\{1, 2, \ldots, p\}$ for some p.

Our next theorem concerns the *Schmidt orthogonalisation process* for constructing orthonormal sequences. If X is a normed space and $x_1, \ldots, x_n \in X$, we denote by $\mathfrak{L}(x_1, \ldots, x_n)$ the subspace (necessarily closed) generated by $x_1, \ldots, x_n$.

THEOREM 1.6.5. *Suppose that $(x_1, x_2, x_3, \ldots)$ is a linearly independent sequence of vectors in a Hilbert space $\mathcal{H}$. Then there is an orthonormal sequence $(y_1, y_2, y_3, \ldots)$ in $\mathcal{H}$ such that $\mathfrak{L}(y_1, \ldots, y_n) = \mathfrak{L}(x_1, \ldots, x_n)$ for each n.*

Proof. We construct $y_1, y_2, \ldots$ inductively, and begin by defining $y_1 = ||x_1||^{-1} x_1$. Now suppose that we have chosen an orthonormal system $(y_1, \ldots, y_{r-1})$ such that

$$\mathfrak{L}(y_1, \ldots, y_n) = \mathfrak{L}(x_1, \ldots, x_n) \quad (1 \leqslant n < r). \tag{5}$$

Since $x_r \notin \mathfrak{L}(x_1, \ldots, x_{r-1}) = \mathfrak{L}(y_1, \ldots, y_{r-1})$, the vector

$$z_r = x_r - \sum_{j=1}^{r-1} \langle x_r, y_j \rangle y_j \tag{6}$$

is non-zero, and is clearly orthogonal to $y_1, \ldots, y_{r-1}$. It follows that, if $y_r = ||z_r||^{-1} z_r$, then $(y_1, \ldots, y_r)$ is an orthonormal system. By (5), each of $y_1, \ldots, y_{r-1}$ is a linear combination of $x_1, \ldots, x_r$, and (6) implies that the same is true of z_r, hence also of y_r. Thus

$$\mathfrak{L}(y_1, \ldots, y_r) \subseteq \mathfrak{L}(x_1, \ldots, x_r);$$

since both these subspaces are r-dimensional, they are equal.

In this way, we construct inductively an orthonormal sequence $(y_1, y_2, \ldots)$ such that (5) is satisfied for all n. If there are just m vectors $x_1, \ldots, x_m$ in the given linearly independent sequence, the process terminates with the construction of y_m; if the sequence (x_n) is infinite, then so is (y_n).

THEOREM 1.6.6. *If M is a closed subspace of a Hilbert space $\mathcal{H}$, then $M \cap M^{\perp} = \{0\}$, $M + M^{\perp} = \mathcal{H}$ and $(M^{\perp})^{\perp} = M$.*

Proof. We have already noted that $B \cap B^{\perp} = \{0\}$ for any non-empty subset B of $\mathcal{H}$. Let A be an orthonormal basis of the

Hilbert space M (hence an orthonormal set in $\mathcal{H}$). Since A generates the closed subspace M, we have $A^{\perp} = M^{\perp}$. Given x in $\mathcal{H}$, the sum $x_1 = \sum_{a \in A} \langle x, a \rangle a$ exists and $x - x_1 \in A^{\perp} = M^{\perp}$, by Lemma 1.6.2. Clearly, $x_1 \in M$, so

$$x = x_1 + (x - x_1) \in M + M^{\perp}.$$

This shows that $M + M^{\perp} = \mathcal{H}$.

It is apparent that $M \subseteq (M^{\perp})^{\perp}$, so $x_1 \in (M^{\perp})^{\perp}$. If, also, $x \in (M^{\perp})^{\perp}$, then $x - x_1 \in (M^{\perp})^{\perp} \cap M^{\perp} = \{0\}$; and thus $x = x_1 \in M$. It follows that $M = (M^{\perp})^{\perp}$.

COROLLARY 1.6.7. *If B is a subset of a Hilbert space $\mathcal{H}$, then $(B^{\perp})^{\perp}$ is the closed subspace M generated by B.*

Proof. $B^{\perp} = M^{\perp}$, so $(B^{\perp})^{\perp} = (M^{\perp})^{\perp} = M$.

THEOREM 1.6.8. *Suppose that $\mathcal{H}$ is a Hilbert space and $y \in \mathcal{H}$. Then the equation $f_y(x) = \langle x, y \rangle$ $(x \in \mathcal{H})$ defines a bounded linear functional f_y on $\mathcal{H}$, and $\|f_y\| = \|y\|$. If f is a bounded linear functional on $\mathcal{H}$, then there is a unique element z of $\mathcal{H}$ such that $f = f_z$.*

Proof. Since $|f_y(x)| = |\langle x, y \rangle| \leqslant \|x\|\ \|y\|$, with equality when $x = y$, it follows that f_y is a bounded linear functional on $\mathcal{H}$, and $\|f_y\| = \|y\|$. To prove that each bounded linear functional f on $\mathcal{H}$ has the form f_z, for some z in $\mathcal{H}$, it suffices to consider the case in which $f \neq 0$. The closed subspace $M = f^{-1}(0)$ is not the whole of $\mathcal{H}$, so $M^{\perp} \neq \{0\}$ by Theorem 1.6.6. Let y be a unit vector in $M^{\perp}$. For each x in $\mathcal{H}$, the vector $f(x)y - f(y)x$ is in M, hence is orthogonal to y. Thus

$$0 = \langle f(x)y - f(y)x,\ y \rangle = f(x) - f(y) \langle x, y \rangle = f(x) - \langle x, z \rangle,$$

where $z = \overline{f(y)}y$. Hence $f = f_z$.

In later chapters we shall need the idea of the direct sum of a family of Hilbert spaces. Suppose that A is a set and, for each a in A, a Hilbert space $\mathcal{H}_a$ is specified. We denote by $\Sigma \oplus \{\mathcal{H}_a : a \in A\}$ the set of all functions f on A such that

$$f(a) \in \mathcal{H}_a \ (a \in A), \quad \sum_{a \in A} \|f(a)\|^2 < \infty.$$

It is not difficult to verify that $\Sigma \oplus \{\mathcal{H}_a : a \in A\}$ becomes a Hilbert space when the algebraic operations and inner product are defined by

$$(\alpha f + \beta g)(a) = \alpha f(a) + \beta g(a),$$
$$\langle f, g \rangle = \sum_{a \in A} \langle f(a), g(a) \rangle.$$

We call this space the *Hilbert direct sum* of the family $\{\mathcal{H}_a : a \in A\}$ (when each $\mathcal{H}_a$ is the complex field, it coincides with $\ell_2(A)$). The Hilbert direct sum of a finite family $\{\mathcal{H}_j : j = 1, \ldots, n\}$ is sometimes denoted by

$$\mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \cdots \oplus \mathcal{H}_n.$$

Given b in A and x in $\mathcal{H}_b$, let $U_b x$ denote the element f of $\Sigma \oplus \{\mathcal{H}_a : a \in A\}$ defined by $f(b) = x$ and $f(a) = 0$ when $a \in A$, $a \neq b$. It is clear that U_b is an isometric isomorphism from $\mathcal{H}_b$ onto a closed subspace of $\Sigma \oplus \{\mathcal{H}_a : a \in A\}$. We frequently identify $\mathcal{H}_b$ with its image $U_b(\mathcal{H}_b)$ under this canonical mapping U_b.

If $\{\mathcal{H}_a : a \in A\}$ is a family of pairwise orthogonal closed subspaces of a Hilbert space $\mathcal{H}$, the sum

$$Uf = \sum_{a \in A} f(a)$$

exists, whenever $f \in \Sigma \oplus \{\mathcal{H}_a : a \in A\}$, by Lemma 1.6.4, and

$$||Uf|| = [\sum_{a \in A} ||f(a)||^2]^{1/2} = ||f||.$$

It follows that U is an isometric isomorphism from $\Sigma \oplus \{\mathcal{H}_a : a \in A\}$ onto a closed subspace M of $\mathcal{H}$, and it is easily verified that M is the smallest closed subspace which contains each $\mathcal{H}_a$. We frequently identify $\Sigma \oplus \{\mathcal{H}_a : a \in A\}$ with M by means of the canonical mapping U.

We conclude this section with another example of a Hilbert space. Suppose that $(E, \mathcal{E}, \mu)$ is a measure space, and that $\mathcal{H}$ is the vector space $L_2(E, \mu)$ of all (equivalence classes of) complex-valued μ-measurable functions f on E for which

$$\int_E |f|^2 \, d\mu < \infty.$$

The equation

$$\langle f, g\rangle = \int_E f\bar{g} \, d\mu$$

defines an inner product on $\mathcal{H}$, and $\mathcal{H}$ is complete with respect to the associated norm,

$$||f|| = \langle f, f\rangle^{1/2} = [\int_E |f|^2 \, d\mu]^{1/2}$$

Thus, $L_2(E, \mu)$ is a Hilbert space. For the proofs of these statements, we refer to [25: p. 174-77].

1.7. *Bounded linear operators on Hilbert space*

Suppose that $\mathcal{H}$ is a Hilbert space and p is a complex-valued function on the Cartesian product $\mathcal{H} \times \mathcal{H}$. We say that p is a *sesqui-linear form on* $\mathcal{H}$ if

$$p(\alpha x+\beta y, z) = \alpha p(x, z)+\beta p(y, z),$$
$$p(x, \alpha y+\beta z) = \bar{\alpha} p(x, y)+\bar{\beta} p(x, z),$$

whenever $x, y, z \in \mathcal{H}$ and α, β are complex numbers. A sesqui-linear form p on $\mathcal{H}$ is *bounded* if there is a real number M such that

$$|p(x, y)| \leqslant M||x|| \, ||y|| \quad (x, y \in \mathcal{H});$$

when this is so, the least such M is denoted by $||p||$.

THEOREM 1.7.1. *If $T \in \mathcal{B}(\mathcal{H})$ and*

(1) $$p_T(x, y) = \langle Tx, y\rangle \quad (x, y \in \mathcal{H}),$$

then p_T is a bounded sesqui-linear form on $\mathcal{H}$, and $||p_T|| = ||T||$. If p is any bounded sesqui-linear form on $\mathcal{H}$, there is a unique element S of $\mathcal{B}(\mathcal{H})$ such that $p = p_S$.

Proof. Clearly p_T is a bounded sesqui-linear form, and $||p_T|| \leqslant ||T||$, for each T in $\mathcal{B}(\mathcal{H})$. Furthermore $||Tx||^2 = \langle Tx, Tx\rangle = p_T(x, Tx) \leqslant ||p_T|| \, ||x|| \, ||Tx||$, and thus $||Tx|| \leqslant ||p_T|| \, ||x||$, for each x in $\mathcal{H}$; so $||p_T|| = ||T||$.

If p is a bounded sesqui-linear form on $\mathcal{H}$, and $x \in \mathcal{H}$, the mapping $y \to \overline{p(x, y)}$ is a bounded linear functional on $\mathcal{H}$, with norm not exceeding $||p|| \, ||x||$. By Theorem 1.6.8 there is a unique element Tx of $\mathcal{H}$ such that $||Tx|| \leqslant ||p|| \, ||x||$ and $\overline{p(x, y)} = \langle y, Tx\rangle$, whence $p(x, y) = \langle Tx, y\rangle$. Since p is linear in the first variable, the mapping $x \to Tx$ is linear. Thus $T \in \mathcal{B}(\mathcal{H})$ and $p_T = p$.

If $T \in \mathcal{B}(\mathcal{H})$ and $u, v \in \mathcal{H}$, a straightforward calculation shows that

(2) $$4\langle Tu, v\rangle = \langle T(u+v), u+v\rangle - \langle T(u-v), u-v\rangle + i\langle T(u+iv), u+iv\rangle - i\langle T(u-iv), u-iv\rangle.$$

LEMMA 1.7.2. *If $T \in \mathcal{B}(\mathcal{H})$ and $\langle Tx, x\rangle = 0$ for each x in $\mathcal{H}$, then $T = 0$.*

Proof. By (2), $p_T(u, v) = \langle Tu, v\rangle = 0$ for each u and v in $\mathcal{H}$. It follows from Theorem 1.7.1 that $||T|| = ||p_T|| = 0$.

If $T \in \mathcal{B}(\mathcal{H})$, the equation $p(x, y) = \langle x, Ty\rangle$ defines a bounded sesqui-linear form p on $\mathcal{H}$. By Theorem 1.7.1, there is a unique element T^* of $\mathcal{B}(\mathcal{H})$ such that $p = p_{T^*}$; that is,

(3) $$\langle x, Ty\rangle = \langle T^*x, y\rangle \quad (x, y \in \mathcal{H}).$$

The operator T^* is called the *adjoint* of T. From (3) and Theorem 1.7.1, it is easily verified that $||T|| = ||T^*||$, and that

$$(A+B)^* = A^*+B^*, \qquad (aA)^* = \bar{a}A^*,$$
$$(AB)^* = B^*A^*, \qquad (A^*)^* = A.$$

It follows that A^* has an inverse if and only if A has an inverse, and then $(A^*)^{-1} = (A^{-1})^*$. By applying this with $\lambda I - T$ in place of A, we conclude that

(4) $$\sigma(T^*) = \{\bar{\lambda} : \lambda \in \sigma(T)\}.$$

Since $||Tx||^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle \leqslant ||T^*T||\,||x||^2$, we have $||T||^2 \leqslant ||T^*T|| \leqslant ||T^*||\,||T|| = ||T||^2$, so

(5) $$||T^*T|| = ||T||^2.$$

With each T in $\mathcal{B}(\mathcal{H})$ we associate two closed subspaces of $\mathcal{H}$, the *null space* $\mathcal{N}_T = \{x \in \mathcal{H} : Tx = 0\}$ of T, and the *closed range space* $\mathcal{R}_T$ of T, which is the closure of the set $\{Tx : x \in \mathcal{H}\}$. From (3), $x \in \mathcal{R}_T^\perp$ if and only if $T^*x = 0$, so

(6) $$\mathcal{R}_T^\perp = \mathcal{N}_{T^*}, \qquad \mathcal{R}_T = (\mathcal{R}_T^\perp)^\perp = \mathcal{N}_{T^*}^\perp.$$

Clearly $\mathcal{N}_T \subseteq \mathcal{N}_{T^*T}$, and the reverse inclusion holds since $||Tx||^2 = \langle T^*Tx, x\rangle$. Thus

(7) $$\mathcal{N}_{T^*T} = \mathcal{N}_T, \qquad \mathcal{R}_{T^*T} = \mathcal{N}_{T^*T}^\perp = \mathcal{N}_T^\perp = \mathcal{R}_{T^*}.$$

An element A of $\mathcal{B}(\mathcal{H})$ is said to be *self-adjoint* if $A = A^*$. Each T in $\mathcal{B}(\mathcal{H})$ can be expressed (uniquely) as $A+iB$, with A and B self-adjoint. The operator A $(= \tfrac{1}{2}(T+T^*))$ is called the *self-adjoint part* (or *real part*) of T, and is denoted by Re T; while B $(= \tfrac{1}{2}i(T^*-T))$ is called the *skew-adjoint part* (or *imaginary part*), Im T. We say that T is *normal* if $T^*T = TT^*$; this occurs if and only if $AB = BA$.

LEMMA 1.7.3. *Suppose that $T \in \mathcal{B}(\mathcal{H})$.*

(i) T is self-adjoint if and only if $\langle Tx, x\rangle$ is real for each x in $\mathcal{H}$.

(ii) T is normal if and only if $||Tx|| = ||T^*x||$ for each x in $\mathcal{H}$.

Proof. Since

$$\langle (T-T^*)x, x\rangle = \langle Tx, x\rangle - \langle x, Tx\rangle = 2i \operatorname{Im} \langle Tx, x\rangle,$$
$$\langle (T^*T-TT^*)x, x\rangle = ||Tx||^2 - ||T^*x||^2$$

for each x in $\mathcal{H}$, the stated results follow at once from Lemma 1.7.2.

THEOREM 1.7.4. *If T is a normal element of $\mathcal{B}(\mathcal{H})$, then $r(T) = ||T||$.*

Proof. By Lemma 1.7.3, $||T^2x|| = ||T^*Tx||$ for each x in $\mathcal{H}$. This, together with (5), implies that

$$||T^2|| = ||T^*T|| = ||T||^2.$$

Since T^n is normal, $||T^{2n}|| = ||T^n||^2$ $(n = 1, 2, \ldots)$. Hence

$$||T^k|| = ||T||^k, \qquad ||T^k||^{1/k} = ||T||$$

when k has the form 2^n $(n = 1, 2, \ldots)$. It follows that

$$r(T) = \lim_{k\to\infty} ||T^k||^{1/k} = ||T||.$$

COROLLARY 1.7.5. *If $T \in \mathcal{B}(\mathcal{H})$ and T is both normal and quasi-nilpotent, then $T = 0$.*

THEOREM 1.7.6. *If T is a self-adjoint element of $\mathcal{B}(\mathcal{H})$, then $\sigma(T)$ is a subset of the real line.*

Proof. We have to show that $T - \lambda I$ has a bounded inverse if $\lambda = \mu + i\nu$ and $\nu \neq 0$. If $x \in \mathcal{H}$, then $\langle Tx, x\rangle$ is real, so

$$\begin{aligned}||Tx-(\mu\pm i\nu)x||\ ||x|| &\geqslant |\langle Tx-(\mu\pm i\nu)x, x\rangle| \\ &= |\langle Tx, x\rangle-(\mu\pm i\nu)||x||^2| \geqslant |\nu|\ ||x||^2;\end{aligned}$$

(8) $$||Tx-(\mu\pm i\nu)x|| \geqslant |\nu|\ ||x|| \quad (x \in \mathcal{H}).$$

In particular, both $T-\lambda I$ and its adjoint $T-\bar{\lambda}I$ have null space $\{0\}$. By (6), the closed range space of $T-\lambda I$ is the whole of $\mathcal{H}$. Thus each y in $\mathcal{H}$ is the limit of a sequence $((T-\lambda I)x_n)$, with $x_1, x_2, \ldots$ in $\mathcal{H}$. By (8),

$$||x_m-x_n|| \leqslant |\nu|^{-1}\ ||(T-\lambda I)(x_m-x_n)|| \to 0$$

as $\min(m, n) \to \infty$. Thus (x_n) converges to some x in $\mathcal{H}$, and $y = \lim(T-\lambda I)x_n = (T-\lambda I)x$. This shows that $T-\lambda I$ maps $\mathcal{H}$ onto the whole of $\mathcal{H}$; by (8), it is one to one, and its inverse R is bounded, with $||R|| \leqslant |\nu|^{-1}$.

If $T = T^* \in \mathcal{B}(\mathcal{H})$, the spectrum $\sigma(T)$ is a compact subset of the real line. The set $C(\sigma(T))$ of all complex-valued continuous functions on $\sigma(T)$ is a Banach space with respect to the uniform norm,

$$||f|| = \sup\{|f(\lambda)| : \lambda \in \sigma(T)\}.$$

When $f, g \in C(\sigma(T))$, the product fg and the complex conjugate function $\bar{f}$ (both defined pointwise) are also in $C(\sigma(T))$. By the classical Weierstrass approximation theorem theorem, the set P of all polynomials (considered as functions on $\sigma(T)$) is an everywhere dense subspace of $C(\sigma(T))$.

THEOREM 1.7.7. *If $T = T^* \in \mathcal{B}(\mathcal{H})$, there is a unique linear mapping $f \to f(T)$ from $C(\sigma(T))$ into $\mathcal{B}(\mathcal{H})$ such that*

(i) *$f(T)$ has its elementary meaning when f is a polynomial;*

(ii) $||f(T)|| = ||f||$ $(f \in C(\sigma(T)))$.

Furthermore, for each f and g in $C(\sigma(T))$,

(iii) $(fg)(T) = f(T)g(T)$;

(iv) $\bar{f}(T) = f(T)^*$;

(v) *$f(T)$ is normal;*

(vi) *$f(T)S = Sf(T)$ whenever $S \in \mathcal{B}(\mathcal{H})$ and $TS = ST$;*

(vii) *if $\lambda \in \sigma_p(T)$, $x \in \mathcal{H}$ and $Tx = \lambda x$, then $f(T)x = f(\lambda)x$.*

Proof. If $p(\lambda) = a_0+a_1\lambda+ \ldots +a_n\lambda^n$, then $\bar{p}(\lambda) = \bar{a}_0+\bar{a}_1\lambda+ \ldots +\bar{a}_n\lambda^n$ when $\lambda \in \sigma(T)$. Thus

$$p(T)^* = (a_0I+a_1T+\dots+a_nT^n)^*$$
$$= \bar{a}_0I+\bar{a}_1T+\dots+\bar{a}_nT^n = \bar{p}(T).$$

It follows that $p(T)$ is normal. By Theorems 1.5.1, 1.5.2 and 1.7.4,

$$||p(T)|| = r(p(T)) = \sup\{|\mu| : \mu \in \sigma(p(T))\}$$
$$= \sup\{|p(\lambda)| : \lambda \in \sigma(T)\} = ||p||,$$

where $||p||$ denotes the norm of p as an element of $C(\sigma(T))$. This shows that the linear mapping $p \to p(T)$ from P into $\mathcal{B}(\mathcal{H})$ is isometric. Since P is dense in $C(\sigma(T))$ and $\mathcal{B}(\mathcal{H})$ is complete, this mapping extends uniquely, by continuity, to an isometric linear mapping $f \to f(T)$ from $C(\sigma(T))$ into $\mathcal{B}(\mathcal{H})$.

We have now proved the existence and uniqueness of a linear mapping $f \to f(T)$ satisfying parts (i) and (ii) of the theorem. It is evident that parts (iii), ..., (vii) are satisfied when f and g are polynomials; by continuity, they remain valid whenever f, $g \in C(\sigma(T))$. This completes the proof of the theorem.

The mapping $f \to f(T)$ described in Theorem 1.7.7 is called the *functional calculus* for the self-adjoint operator T. A further property of this mapping is proved in Theorem 1.7.9.

Suppose that S and T are self-adjoint elements of $\mathcal{B}(\mathcal{H})$. We write $S \leqslant T$ if $\langle Sx, x\rangle \leqslant \langle Tx, x\rangle$ for each x in $\mathcal{H}$. If $S \leqslant T$, and $S \neq T$, we write $S < T$. It follows from Lemma 1.7.2 that $\leqslant$ is a partial order relation. It is easily verified that, if $S \leqslant T$, then $A^*SA \leqslant A^*TA$ for each A in $\mathcal{B}(\mathcal{H})$. Furthermore, $S \leqslant ||S||I$ for every self-adjoint S in $\mathcal{B}(\mathcal{H})$.

If $T \in \mathcal{B}(\mathcal{H})$ and $\langle Tx, x\rangle$ is a non-negative real number whenever $x \in \mathcal{H}$, it follows from Lemma 1.7.3 that T is self-adjoint,

and so $T \geqslant 0$; in these circumstances, we call T a *positive operator.* If S and T are positive operators and α is a positive real number, then $S+T \geqslant 0$, $\alpha S \geqslant 0$.

THEOREM 1.7.8. *If T is a self-adjoint element of $\mathcal{B}(\mathcal{H})$, then the following conditions are equivalent.*

(i) $\sigma(T)$ consists of non-negative real numbers.

(ii) $T = H^2$ for some positive operator H.

(iii) $T = A^*A$ for some A in $\mathcal{B}(\mathcal{H})$.

(iv) $T \geqslant 0$.

When these conditions are satisfied, the operator H occuring in (ii) is unique.

Proof. It is clear that (ii) implies (iii). If $T = A^*A$, then $\langle Tx, x\rangle = \langle Ax, Ax\rangle \geqslant 0$, for each x in $\mathcal{H}$; so (iii) implies (iv).

If (i) is satisfied, the real-valued function $g(\lambda) = \lambda^{1/4}$ is continuous on $\sigma(T)$, and $[g(\lambda)]^4 = \lambda$. Hence the operator $S = g(T)$ is self-adjoint, and $S^4 = T$. Since (iii) implies (iv), $H = S^2$ is a positive operator and $H^2 = T$. Thus (i) implies (ii).

Suppose $T \geqslant 0$, and define h in $C(\sigma(T))$ by

$$h(\lambda) = \begin{cases} 0 & \text{when } \lambda \geqslant 0, \\ (-\lambda)^{1/2} & \text{when } \lambda < 0. \end{cases} \qquad (9)$$

Since h is real-valued and $h(\lambda)\lambda h(\lambda) = -[h(\lambda)]^4$ when $\lambda \in \sigma(T)$, the operator $S = h(T)$ is self-adjoint and $STS = -S^4$. Since (iii) implies (iv) we have $S^4 \geqslant 0$, and so

$$0 \geqslant -S^4 = STS = STS^* \geqslant 0.$$

Thus $0 = S^4 = [h(T)]^4$ and, since the mapping $f \to f(T)$ is one to one, it follows that h^4 is the zero element of $C(\sigma(T))$. This, together

with (9), shows that $\sigma(T)$ consists of non-negative real numbers. Thus (iv) implies (i), and the equivalence of the four conditions in the theorem is proved.

It remains to show that, if $T \geqslant 0$, the operator H in (iii) is unique. Now $H = f(T)$, where f is the element of $C(\sigma(T))$ defined by $f(\lambda) = \lambda^{1/2}$. If (p_n) is a sequence of polynomials converging uniformly to f on $\sigma(T)$, then

$$(10) \qquad ||H-p_n(T)|| \to 0 \quad \text{as } n \to \infty.$$

Suppose that K is any positive operator such that $K^2 = T$. Since $\sigma(K)$ consists of non-negative real numbers, and $\sigma(T) = \{\mu^2 : \mu \in \sigma(K)\}$, the polynomial $q_n(\mu) = p_n(\mu^2)$ converges, uniformly on $\sigma(K)$, to the function $f(\mu^2) = \mu = u(\mu)$, where u is the identity mapping on $\sigma(K)$. Thus

$$(11) \qquad ||K-p_n(T)|| = ||K-q_n(K)|| = ||u-q_n|| \to 0 \quad \text{as } n \to \infty.$$

By (10) and (11), $H = K$. This shows that H is unique, and so completes the proof of the theorem.

When $T \geqslant 0$, the operator H in Theorem 1.7.8(ii) is called the *positive square root* of T, and is denoted by $T^{1/2}$. Note that, if $T \geqslant 0$, $x \in \mathcal{H}$ and $\langle Tx, x\rangle = 0$, then $Tx = 0$; for

$$||T^{1/2}x||^2 = \langle T^{1/2}x, T^{1/2}x\rangle = \langle Tx, x\rangle = 0,$$

and so $Tx = T^{1/2}T^{1/2}x = 0$.

THEOREM 1.7.9. *Suppose that T is a self-adjoint element of $\mathcal{B}(\mathcal{H})$ and $f \in C(\sigma(T))$. Then $f(T) \geqslant 0$ if and only if f takes non-negative real values throughout $\sigma(T)$.*

Proof. If $f(\lambda) \geqslant 0$ $(\lambda \in \sigma(T))$, then there is a real-valued function g in $C(\sigma(T))$ satisfying $g^2 = f$. The operator $A = g(T)$ is self-adjoint and $f(T) = A^2$; so $f(T) \geqslant 0$.

Conversely, suppose that $f(T) \geqslant 0$. Since $f(T)$ is self-adjoint, f is real-valued on $\sigma(T)$; and so, $f = u-v$, where

$$u(\lambda) = \max\{0, f(\lambda)\}, \qquad v(\lambda) = \max\{0, -f(\lambda)\}.$$

Since v^3 is a non-negative function, in $C(\sigma(T))$, it follows from the first part of this proof that $[v(T)]^3 \geqslant 0$. However, $vfv = -v^3$, so

$$0 \leqslant v(T)^*f(T)v(T) = -[v(T)]^3 \leqslant 0.$$

Thus $[v(T)]^3 = 0$; it follows that v^3, and hence also v, is the zero element of $C(\sigma(T))$. Thus $f(\lambda) = u(\lambda) \geqslant 0$ for each λ in $\sigma(T)$.

We now describe the simplest class of self-adjoint operators. If M is a closed subspace of a Hilbert space $\mathcal{H}$, then $M \cap M^\perp = \{0\}$, $M + M^\perp = \mathcal{H}$. By elementary linear algebra (see, for example [26: pp. 73, 74]) there is a unique linear operator E on $\mathcal{H}$ such that $Ex \in M$ and $x-Ex \in M^\perp$ for each x in $\mathcal{H}$. Furthermore, $E^2 = E$ and

$$(12) \qquad M = \{x \in \mathcal{H} : Ex = x\} = \{Ex : x \in \mathcal{H}\},$$

$$(13) \qquad M^\perp = \{x \in \mathcal{H} : Ex = 0\}.$$

Since Ex and $x-Ex$ are orthogonal, with sum x, we have

$$||x||^2 = ||Ex||^2 + ||x-Ex||^2 \geqslant ||Ex||^2 \quad (x \in \mathcal{H}).$$

Thus E is a bounded linear operator, with $||E|| \leqslant 1$. Furthermore, $\langle Ex, x-Ex\rangle = 0$, and therefore

$$\langle Ex, x\rangle = \langle Ex, Ex\rangle = ||Ex||^2 \geqslant 0 \quad (x \in \mathcal{H}).$$

It follows that E is a positive self-adjoint operator. We call E the *projection from* $\mathcal{H}$ *onto* M (it is sometimes described as an *orthogonal* projection, since its range space M is orthogonal to its null space $M^\perp$).

Conversely, if $E \in \mathcal{B}(\mathcal{H})$ and $E^2 = E = E^*$, then (12) defines a closed subspace M of $\mathcal{H}$. Since $M = \mathcal{R}_E$, it follows from (6) that $M^\perp = \mathcal{N}_{E^*} = \mathcal{N}_E$, so (13) is satisfied. If $x \in \mathcal{H}$ then clearly $Ex \in M$; moreover, $x-Ex \in M^\perp$ by (13), since $E(x-Ex) = Ex-E^2x = 0$. Hence E is the projection from $\mathcal{H}$ onto M. It follows that projections are characterized by the conditions $E^2 = E = E^*$.

Suppose that E and F are the projections from $\mathcal{H}$ onto closed subspaces M and N, respectively. It is readily verified that M is orthogonal to N if and only if $EF = 0$ (equivalently, since $FE = (EF)^*$, $FE = 0$). If $FE = EF$, then each of the operators EF, $E+F-EF$ satisfies the equations $P^2 = P^* = P$, and is therefore a projection; a straightforward argument shows that the corresponding closed subspaces of $\mathcal{H}$ are $M \cap N$ and $M+N$, respectively.

Elementary linear algebra (see, for example, [26: p. 148]) suffices to prove the following result.

LEMMA 1.7.10. *If E and F are the projections from a Hilbert space $\mathcal{H}$ onto closed subspaces M and N, respectively, then the following conditions are equivalent.*

(i) $M \subseteq N$.
(ii) $FE = E$.
(iii) $EF = E$.
(iv) $||Ex|| \leqslant ||Fx|| \quad (x \in \mathcal{H})$.
(v) $E \leqslant F$.

When these conditions are satisfied, $F-E$ is the projection from $\mathcal{H}$ onto $N \cap M^\perp$. In particular, $I-E$ is the projection from $\mathcal{H}$ onto $M^\perp$.

If (M_α) is a family of closed subspaces of a Hilbert space $\mathcal{H}$, there is a smallest closed subspace $\vee M_\alpha$ which contains each M_α, and a largest closed subspace $\wedge M_\alpha$ $(= \cap M_\alpha)$ contained in each M_α. Lemma 1.7.10 shows that the set of closed subspaces of $\mathcal{H}$ (partially ordered by $\subseteq$) is order isomorphic to the set $\mathcal{P}$ of projections (partially ordered by $\leqslant$). Thus each family (E_α) of projections has a least upper bound $\vee E_\alpha$ and a greatest lower bound $\wedge E_\alpha$ in $\mathcal{P}$. We use the notations $M \vee N$, $M \wedge N$, $E \vee F$, $E \wedge F$ when only two subspaces or projections are under consideration. A family $\{E_\alpha : \alpha \in A\}$ of projections is said to be *directed upward* if, given any β and γ in A, there is a δ in A such that $E_\delta \geqslant E_\beta$, $E_\delta \geqslant E_\gamma$. The definition of *directed downward* is similar, but $\geqslant$ is replaced by $\leqslant$.

The *strong topology* on $\mathcal{B}(\mathcal{H})$ is the coarsest topology such that, for each x in $\mathcal{H}$, the mapping $T \to Tx$ from $\mathcal{B}(\mathcal{H})$ into $\mathcal{H}$ is continuous. If $E_0 \in \mathcal{B}(\mathcal{H})$, then sets of the form

$$(14) \qquad \{E \in \mathcal{B}(\mathcal{H}) : ||Ex_j - E_0x_j|| < \epsilon \quad (j = 1, \ldots, n)\}$$

(where $x_1, \ldots, x_n \in \mathcal{H}$ and $\epsilon > 0$) form a base of neighbourhoods of E_0 in the strong topology. In the following theorem, 'strong closure' is used as an abbreviation for 'closure in the strong topology'. We shall adopt similar conventions when using other topological terms in relation to the strong topology.

THEOREM 1.7.11. *Suppose that $\{E_\alpha : \alpha \in A\}$ is a family of projections acting on a Hilbert space $\mathcal{H}$, and is directed upward (respectively, downward). Then $\vee E_\alpha$ (respectively, $\wedge E_\alpha$) lies in the strong closure of $\{E_\alpha : \alpha \in A\}$.*

Proof. We consider first the case in which $\{E_\alpha : \alpha \in A\}$ is directed upward. Suppose that E_α is the projection from $\mathcal{H}$ onto the closed subspace M_α, so that $E_0 = \vee E_\alpha$ is the projection onto the closed subspace M_0 generated by $\cup M_\alpha$. In fact, $\cup M_\alpha$ is already a subspace, and so M_0 is its closure: for, if $x, y \in \cup M_\alpha$, we may choose β and γ, and then δ, such that $x \in M_\beta$, $y \in M_\gamma$, $M_\beta \subseteq M_\delta$, $M_\gamma \subseteq M_\delta$; it follows that $x, y \in M_\delta$, and so $z \in M_\delta \subseteq \cup M_\alpha$, whenever z is a linear combination of x and y.

Every strong neighbourhood of E_0 contains a set of the form (14), so it suffices to show that this set contains an E_α. Since E_0x_j lies in the closure M_0 of $\cup M_\alpha$, there exist $\alpha(j)$ in A and y_j in $M_{\alpha(j)}$ such that $||E_0x_j-y_j|| < \epsilon$. There is an element β of A such that $M_{\alpha(j)} \subseteq M_\beta$ $(j = 1, \ldots, n)$, and so $y_1, \ldots, y_n \in M_\beta$. Thus

$$\begin{aligned} ||E_0x_j-E_\beta x_j|| &= ||(E_0-E_\beta)(E_0x_j-y_j)|| \\ &\leqslant ||E_0x_j-y_j|| < \epsilon \quad (j = 1, \ldots, n). \end{aligned}$$

Hence E_β lies in the set (14). This proves the lemma in the case where $\{E_\alpha : \alpha \in A\}$ is directed upward.

Now suppose that $\{E_\alpha : \alpha \in A\}$ is directed downward, and let $E_0 = \wedge E_\alpha$. Since two projections P, Q satisfy $P \leqslant Q$ if and only if $I-P \geqslant I-Q$, it follows that $\{I-E_\alpha : \alpha \in A\}$ is directed upward and $\vee(I-E_\alpha) = I-E_0$. It follows from the first part of the proof that $I-E_0$ lies in the strong closure of $\{I-E_\alpha : \alpha \in A\}$. If $x_1, \ldots, x_n \in \mathcal{H}$ and $\epsilon > 0$, there exists β in A such that

$$||(I-E_\beta)x_j-(I-E_0)x_j|| < \epsilon \quad (j = 1, \ldots, n).$$

Thus E_β lies in the basic neighbourhood (14) of E_0; and so E_0 is in the strong closure of $\{E_\alpha : \alpha \in A\}$.

Suppose that $\mathcal{H}$ is a Hilbert space and $V \in \mathcal{B}(\mathcal{H})$. We say that V is a *partial isometry* if there is a closed subspace M of $\mathcal{H}$ such that

$$(15) \qquad ||Vx|| = ||x|| \quad (x \in M), \qquad Vx = 0 \quad (x \in M^\perp).$$

When this condition is satisfied, the restriction of V to M is an isometric isomorphism from M onto the subspace $N = \{Vx : x \in M\}$ of $\mathcal{H}$. Since M is a closed subspace of $\mathcal{H}$, it is complete; thus N is complete, and is therefore closed in $\mathcal{H}$. Since $M+M^\perp = \mathcal{H}$, it follows from (15) that

$$(16) \qquad N = \{Vx : x \in M\} = \{Vx : x \in \mathcal{H}\}.$$

We call M the *initial space* and N the *final space* of V; the corresponding projections E and F are called the *initial* and *final projections* of V. Clearly

$$(17) \qquad V = FVE.$$

If $x \in \mathcal{H}$, then $Ex \in M$ and, by (15),

$$\begin{aligned} \langle V^*Vx, x\rangle &= \langle Vx, Vx\rangle = \langle VEx, VEx\rangle \\ &= \langle Ex, Ex\rangle = \langle Ex, x\rangle. \end{aligned}$$

By Lemma 1.7.2,

$$(18) \qquad V^*V = E.$$

Given x in $\mathcal{H}$, it follows from (16) that $Fx = Vy$ for some y in $\mathcal{H}$. By (17), $V^* = V^*F$ and $V = VE$, so

$$VV^*x = VV^*Fx = VV^*Vy = VEy = Vy = Fx.$$

Thus

$$(19) \qquad VV^* = F.$$

If $x \in N$, then

$$||V^*x||^2 = \langle VV^*x, x\rangle = \langle Fx, x\rangle = ||x||^2;$$

if $x \in N^\perp$, then $Fx = 0$ and, by (17), $V^*x = V^*Fx = 0$. Thus V^* is a partial isometry, with initial projection F; by (18) and (19), its final projection is E.

By a *unitary operator* on a Hilbert space $\mathcal{H}$ we mean an invertible element U of $\mathcal{B}(\mathcal{H})$ satisfying $||Ux|| = ||x||$ $(x \in \mathcal{H})$. Since U is a partial isometry with initial projection I, it follows that $U^*U = I$. Hence the inverse U^{-1} is U^*, and

$$\text{(20)} \qquad \langle Ux, Uy\rangle = \langle U^*Ux, y\rangle = \langle x, y\rangle \quad (x, y \in \mathcal{H}).$$

Conversely, if U is an invertible element of $\mathcal{B}(\mathcal{H})$ and $U^{-1} = U^*$, then (20) is satisfied; in particular, $||Ux|| = ||x||$ $(x \in \mathcal{H})$, so U is unitary.

THEOREM 1.7.12. *Suppose that $T \in \mathcal{B}(\mathcal{H})$, and let M and N be the closed range spaces of T^* and T, respectively. Then there is a partial isometry V with initial space M and final space N, and a positive operator H, such that $T = VH$, $H = V^*T$. These conditions determine V and H uniquely.*

Proof. Let H be the positive square root of T^*T. By two applications of (7), the operators H, $H^2 = T^*T$ and T^* all have the same closed range space M. For each x in $\mathcal{H}$,

$$||Tx||^2 = \langle T^*Tx, x\rangle = \langle H^2x, x\rangle = ||Hx||^2.$$

It follows that the mapping $Hx \to Tx$ is well-defined, and is an isometric isomorphism from a dense subspace of M onto a dense subspace of N. This mapping extends by continuity to an isometric isomorphism W from M onto N, such that $WHx = Tx$ $(x \in \mathcal{H})$.

The linear operator V defined by

$$Vy = Wy \quad (y \in M), \qquad Vz = 0 \quad (z \in M^\perp)$$

is a partial isometry with initial space M and final space N; and $Tx = VHx$ $(x \in \mathcal{H})$. Since V^*V is the projection onto M, the closed range space of H, we have

$$Hx = V^*VHx = V^*Tx \quad (x \in \mathcal{H}).$$

This proves the existence of operators V and H with the stated properties.

Suppose that V_1 and H_1 also have these properties. Then $V_1V_1^*$ is the projection onto the closed range space N of T, and so

$$H_1^2 = H_1^*H_1 = T^*V_1V_1^*T = T^*T.$$

Since the positive square root of T^*T is unique, $H_1 = H$. For each x in $\mathcal{H}$,

$$VHx = Tx = V_1H_1x = V_1Hx.$$

Since V and V_1 are continuous, it follows that $Vy = V_1y$ for each y in the closed range space M of H. Furthermore, $Vz = V_1z = 0$ whenever $z \in M^\perp$. Thus $V = V_1$. This shows that the operators V and H are uniquely determined.

When operators T, V and H are related as in Theorem 1.7.12, we describe VH as the *polar decomposition* of T.

1.8. *Compact linear operators*

Suppose that X is a complex Banach space and $T \in \mathcal{B}(X)$. We say that T is a *compact* linear operator on X if, given any sequence (x_n) in X such that $\sup\{||x_n|| : n = 1, 2, \ldots\} < \infty$, there is a subsequence

$(x_{n(q)})$ such that $(Tx_{n(q)})$ converges to an element of X. It is readily verified that the set $\mathcal{C}$ of all compact linear operators on X is an ideal in $\mathcal{B}(X)$; that is RS, SR, $\alpha S+\beta T \in \mathcal{C}$ whenever α, β are complex numbers, $R \in \mathcal{B}(X)$ and S, $T \in \mathcal{C}$. Furthermore, $\mathcal{C}$ is closed with respect to the norm topology of $\mathcal{B}(X)$ [59: p. 285, exercise 3].

An element T of $\mathcal{B}(X)$ is said to have *finite rank* if the subspace $M = \{Tx : x \in X\}$ of X is finite-dimensional; the dimension n of M is then called the *rank* of T. Such an operator T is compact, since every bounded sequence in the finite dimensional normed space M has a convergent subsequence [59: Theorem 3.12-D]. It is readily verified that the operators of finite rank form an ideal in $\mathcal{B}(X)$.

We now turn to the spectral theory of compact linear operators. A full account of this subject is given in [59: pp. 268–85]; accordingly, a number of the results that follow are stated without proof. In the next theorem, T is a compact linear operator acting on a complex Banach space X, λ is a non-zero complex number, and the subspaces $\mathcal{N}_k$ and $\mathcal{R}_k$ of X are defined, for $k = 0, 1, 2, \ldots$, by

$$\mathcal{N}_k = \{x \in X : (\lambda I-T)^k x = 0\}, \tag{1}$$

$$\mathcal{R}_k = \{(\lambda I-T)^k x : x \in X\}. \tag{2}$$

THEOREM 1.8.1. *For $k = 0, 1, 2, \ldots$, the subspaces $\mathcal{R}_k$ and $\mathcal{N}_k$ are invariant under T; $\mathcal{N}_k$ is finite-dimensional and $\mathcal{R}_k$ is closed. There is an integer $\nu(\geqslant 0)$ such that*

(i) $\mathcal{N}_k \subsetneqq \mathcal{N}_{k+1}$ *if* $0 \leqslant k < \nu$, $\mathcal{N}_k = \mathcal{N}_{k+1}$ *if* $k \geqslant \nu$;

(ii) $\mathcal{R}_k \supsetneqq \mathcal{R}_{k+1}$ *if* $0 \leqslant k < \nu$, $\mathcal{R}_k = \mathcal{R}_{k+1}$ *if* $k \geqslant \nu$.

Furthermore,

$$\mathcal{N}_\nu \cap \mathcal{R}_\nu = \{0\}, \qquad \mathcal{N}_\nu + \mathcal{R}_\nu = X.$$

The restriction of $\lambda I-T$ to $\mathcal{R}_\nu$ has an inverse in $\mathcal{B}(\mathcal{R}_\nu)$, and the restriction to $\mathcal{N}_\nu$ of $(\lambda I-T)^\nu$ is 0.

The integer ν occurring in Theorem 1.8.1 is called the *index* of λ with respect to T. If λ is not an eigenvalue of T, then the operator $\lambda I-T$ is one to one, and so $\mathcal{N}_1 = \{0\} = \mathcal{N}_0$. Thus $\nu = 0$, $\mathcal{R}_\nu = \mathcal{R}_0 = X$, and $\lambda I-T$ has an inverse in $\mathcal{B}(X)$. Hence $\lambda \in \rho(T)$. If, on the other hand, λ is an eigenvalue of T, then $\mathcal{N}_1 \neq \{0\} = \mathcal{N}_0$, and so $\nu > 0$. In this case, the dimensions of the subspaces $\mathcal{N}_1$ and $\mathcal{N}_\nu$ are called the *geometric multiplicity* and *algebraic multiplicity* (respectively) of λ as an eigenvalue of T. Note that $\mathcal{N}_1$ consists of all eigenvectors of T associated with the eigenvalue λ, while $\mathcal{N}_\nu$ consists of all principal vectors of T associated with λ.

THEOREM 1.8.2. *Let T be a compact linear operator acting on a Banach space X. If $\lambda \in \sigma(T)$ and $\lambda \neq 0$, then λ is an eigenvalue of T. Furthermore, $\sigma(T)$ is either finite or countably infinite; in the latter case, $\sigma(T) = \{\lambda_n : n = 1, 2, \ldots\}$, where (λ_n) is a complex sequence which converges to 0.*

THEOREM 1.8.3. *Suppose that T is a compact linear operator acting on a complex Banach space X, and M is a closed invariant subspace of T.*

(i) The restriction T^M of T to M is a compact linear operator on M.

(ii) The mapping $T_M : x+M \to Tx+M$ is a compact linear operator on the quotient space X/M.

(iii) If $\lambda \in \rho(T)$ then M is invariant under $(\lambda I-T)^{-1}$.

Proof. Part (i) of the theorem is an immediate consequence of the definition of 'compact linear operator'. It follows from Theorem

1.8.2 that the resolvent set $\rho(T)$ is connected, so (iii) is a consequence of Theorem 1.5.3.

For x in X, let $[x]$ denote the coset $x+M$. Given a sequence $([x_n])$ in X/M such that $\sup \|[x_n]\| < \infty$, we can choose y_n in $[x_n]$ such that $\|y_n\| \leqslant 2\|[x_n]\|$. Thus $\sup \|y_n\| < \infty$, and so there is a subsequence $(y_{n(q)})$ such that $(Ty_{n(q)})$ converges to an element y of X. It follows that

$$\begin{aligned} \|T_M[x_{n(q)}]-[y]\| &= \|T_M[y_{n(q)}]-[y]\| \\ &= \|[Ty_{n(q)}-y]\| < \|Ty_{n(q)}-y\| \to 0 \end{aligned}$$

as $q \to \infty$. This shows that T_M is a compact linear operator on X/M, and completes the proof of the theorem.

We now specialize to the case of a compact linear operator acting on a Hilbert space. The next few theorems can be formulated to as to apply to compact operators acting on Banach spaces (see, for example, [59: pp. 275, 282, 283]), but it is the Hilbert space versions that will be needed in later chapters. The more elementary results are, again, stated without proof.

If T is a compact linear operator acting on a Hilbert space $\mathcal{H}$, then the adjoint T^* is compact. If λ is a non-zero complex number then, in addition to the subspaces $\mathcal{N}_k$ and $\mathcal{R}_k$ introduced in (1) and (2), we define

$$\mathcal{N}_k^* = \{x \in \mathcal{H} : (\bar{\lambda}I-T^*)^k x = 0\}, \tag{3}$$

$$\mathcal{R}_k^* = \{(\bar{\lambda}I-T^*)^k x : x \in \mathcal{H}\}. \tag{4}$$

Since $\mathcal{R}_k$ and $\mathcal{R}_k^*$ are closed (Theorem 1.8.1), they are the closed range spaces of the operator $(\lambda I-T)^k$ and its adjoint $(\bar{\lambda}I-T^*)^k$, respectively. By 1.7(6),

$$\mathcal{R}_k^* = \mathcal{N}_k^\perp, \qquad \mathcal{R}_k = (\mathcal{N}_k^*)^\perp. \tag{5}$$

We use the notation just introduced in the following theorem.

THEOREM 1.8.4. *The spaces $\mathcal{N}_k$ and $\mathcal{N}_k^*$ have the same finite dimension ($k = 0, 1, 2, \ldots$). The index of λ with respect to T is the same as that of $\bar{\lambda}$ with respect to T^*; when it is positive, the geometric and algebraic multiplicities of λ as an eigenvalue of T are the same as those of $\bar{\lambda}$ as an eigenvalue of T^**

It is implicit in Theorem 1.8.4 that a non-zero complex number λ is an eigenvalue of T if and only if $\bar{\lambda}$ is an eigenvalue of T^*; this can also be deduced from Theorem 1.8.2, together with 1.7.(4).

THEOREM 1.8.5. *Suppose that T is a compact linear operator acting on a Hilbert space $\mathcal{H}$, and $\mathcal{P}_T$ is the closed subspace of $\mathcal{H}$ generated by the set of all principal vectors of T associated with non-zero eigenvalues. Then $\mathcal{P}_T^\perp$ is invariant under T^*, and the restriction N of T^* to $\mathcal{P}_T^\perp$ is quasi-nilpotent.*

Proof. If λ is a non-zero eigenvalue of T, let

$$\mathcal{N}(\lambda) = \{x \in \mathcal{H} : (\lambda I-T)^n x = 0 \text{ for some positive integer } n\}, \tag{6}$$

the space of all principal vectors of T associated with λ. Thus $\mathcal{P}_T$ is the closed subspace of $\mathcal{H}$ generated by

$$\cup \{\mathcal{N}(\lambda) : \lambda \in \sigma(T),\ \lambda \neq 0\}.$$

Since each $\mathcal{N}(\lambda)$ is invariant under T, so is $\mathcal{P}_T$. If $y \in \mathcal{P}_T^\perp$ then, for each x in $\mathcal{P}_T$ we have $Tx \in \mathcal{P}_T$ and so

$$\langle T^*y, x\rangle = \langle y, Tx\rangle = 0.$$

Hence $T^*y \in \mathcal{P}_T^\perp$ whenever $y \in \mathcal{P}_T^\perp$; that is, $\mathcal{P}_T^\perp$ is invariant under T^*.

The restriction N of T^* to $\mathcal{P}_T^\perp$ is a compact linear operator on $\mathcal{P}_T^\perp$. Suppose that it is not quasi-nilpotent. Then there is a non-zero μ in $\sigma(N)$ and, by Theorem 1.8.2, μ is an eigenvalue of N. Let x be a non-zero vector in $\mathcal{P}_T^\perp$ such that $Nx = \mu x$ (that is, $T^*x = \mu x$). Since μ is a non-zero eigenvalue of T^*, its complex conjugate λ is a non-zero eigenvalue of T. Let ν be the index of λ relative to T, so that $\bar\lambda$ has index ν relative to T^*. With the notation introduced in (1), (2), (3), (4) and (6), we have $\mathcal{N}_\nu = \mathcal{N}(\lambda) \subseteq \mathcal{P}_T$ and, by (5),

$$\mathcal{P}_T^\perp \subseteq \mathcal{N}_\nu^\perp = \mathcal{R}_\nu^*.$$

It follows from Theorem 1.8.1 that the restriction of $\bar\lambda I - T^*$ to $\mathcal{R}_\nu^*$ has an inverse in $\mathcal{B}(\mathcal{R}_\nu^*)$, and is therefore one to one. However,

$$x \in \mathcal{P}_T^\perp \subseteq \mathcal{R}_\nu^*, \qquad x \neq 0, \qquad (\bar\lambda I - T^*)x = \mu x - T^*x = 0.$$

This contradiction arose from the assumption that N is not quasi-nilpotent; so the theorem is proved.

Suppose that T is a compact linear operator acting on a Hilbert space $\mathcal{H}$. By Theorem 1.8.2, the non-zero elements of $\sigma(T)$ can be arranged as a sequence $(\mu_1, \mu_2, \ldots)$. Suppose that μ_j has algebraic multiplicity $m(j)$ as an eigenvalue of T, and that (λ_n) is the sequence obtained by taking $m(1)$ terms μ_1, then $m(2)$ terms μ_2, then $m(3)$ terms μ_3, and so on. We shall describe (λ_n) as the sequence of non-zero eigenvalues of T, *counted according to their algebraic multiplicities*. The above conditions do not determine (λ_n) uniquely, since the non-zero points of $\sigma(T)$ can be arranged in any order as a sequence (μ_n); usually, this ambiguity does not matter, but occasionally we impose additional conditions on (μ_n) to ensure uniqueness.

THEOREM 1.8.6. *Suppose that T is a compact linear operator acting on a Hilbert space $\mathcal{H}$, (λ_n) is the sequence of non-zero eigenvalues of T (counted according to their algebraic multiplicities) and $\mathcal{P}_T$ is the closed subspace of $\mathcal{H}$ generated by the set of all principal vectors of T associated with the non-zero eigenvalues. Then there is an orthonormal basis $\{\varphi_n : n = 1, 2, \ldots\}$ of $\mathcal{P}_T$ such that $\langle T\varphi_n, \varphi_n\rangle = \lambda_n$.*

Proof. We use the notation introduced in the paragraph preceding the statement of Theorem 1.8.6, and denote by $\nu(j)$ the index of μ_j relative to T. Thus

$$\mathcal{N}(j) = \{x \in \mathcal{H} : (\mu_j I - T)^{\nu(j)}x = 0\} \tag{7}$$

is the space of all principal vectors of T associated with the eigenvalue μ_j; it has dimension $m(j)$, and is invariant under T. If T_j denotes the restriction of T to $\mathcal{N}(j)$, then the operator $\mu_j I - T_j$ on $\mathcal{N}(j)$ is nilpotent. There is a basis $e(j, 1), e(j, 2), \ldots, e(j, m(j))$ of $\mathcal{N}(j)$, with respect to which $\mu_j I - T_j$ is represented by a super-diagonal matrix $A = (a_{k,\ell})$ [26: p. 107]. Since A is nilpotent, its diagonal coefficients (which are its eigenvalues) are all zero. Thus $a_{k,\ell} = 0$ $(k \geq \ell)$, and

$$(\mu_j I - T)e(j, \ell) = \sum_{k=1}^{\ell-1} a_{k,\ell}\, e(j, k). \tag{8}$$

Let (ψ_n) be the sequence of vectors obtained by taking first $e(1, 1), \ldots, e(1, m(1))$, then $e(2, 1), \ldots, e(2, m(2))$, then $e(3, 1), \ldots, e(3, m(3))$, and so on. Since the sequence (λ_n) consists of μ_1 ($m(1)$ times), μ_2($m(2)$ times), and so on, it follows that the integers n for which $\lambda_n = \mu_j$ for some preassigned j are precisely those for which ψ_n has the form $e(j, \ell)$ $(1 \leq \ell \leq m(j))$.

This, together with (8), implies that $(\lambda_n I - T)\psi_n$ can be expressed in the form

(9) $$(\lambda_n I - T)\psi_n = \sum_{k=1}^{n-1} b_{k,n}\psi_k \quad (n = 1, 2, \ldots).$$

We shall now show that the sequence (ψ_n) is linearly independent Suppose that this is not so. Then there exist an integer n and scalars $\alpha(j, k)$ $(k = 1, \ldots, m(j);\ j = 1, \ldots, n)$, not all zero, such that

$$\sum_{j=1}^{n} \sum_{k=1}^{m(j)} \alpha(j, k)\, e(j, k) = 0.$$

Hence

(10) $$x_j \in \mathfrak{N}(j), \ \sum_{j=1}^{n} x_j = 0,$$

where

$$x_j = \sum_{k=1}^{m(j)} \alpha(j, k)\, e(j, k) \quad (j = 1, \ldots, n).$$

Since $e(j, 1), \ldots, e(j, m(j))$ are linearly independent, and some $\alpha(j, k)$ is non-zero, the corresponding x_j is non-zero.

Let $S_j = (\mu_j I - T)^{\nu(j)}$. By (7), $S_j x = 0$ whenever $x \in \mathfrak{N}(j)$. If $\ell \neq j$ and $x \in \mathfrak{N}(\ell)$, then

$$S_j x = (\mu_j I - T)^{\nu(j)} x = (\mu_j I - T_\ell)^{\nu(j)} x.$$

However, since $\mu_\ell I - T_\ell$ is nilpotent, the spectrum of T_ℓ consists of μ_ℓ only, and so $\mu_j I - T_\ell$ is invertible. It follows that

$$S_j x = 0 \ (x \in \mathfrak{N}(j)); \quad S_j x \in \mathfrak{N}(\ell) \sim \{0\} \text{ if } x \in \mathfrak{N}(\ell) \sim \{0\}.$$

Hence the operator $R_\ell = S_1 \ldots S_{\ell-1} S_{\ell+1} \ldots S_n$ annihilates $\mathfrak{N}(j)$ when $1 \leq j \leq n$ and $j \neq \ell$, but is one to one on $\mathfrak{N}(\ell)$. From (10),

$$0 = \sum_{j=1}^{n} R_\ell x_j = R_\ell x_\ell,$$

and so $x_\ell = 0$ $(\ell = 1, \ldots, n)$. This contradicts our previous conclusion that at least one x_j is non-zero, and so establishes our assertion that the sequence (ψ_n) is linearly independent.

Let (φ_n) be the orthonormal sequence obtained from (ψ_n) by the Schmidt orthogonalisation process. With the notation used in Theorem 1.6.5,

(11) $$\varphi_n \in \mathfrak{L}(\varphi_1, \ldots, \varphi_n) = \mathfrak{L}(\psi_1, \ldots, \psi_n).$$

It follows from (9) that $(\lambda_n I - T)\psi_n \in \mathfrak{L}(\psi_1, \ldots, \psi_{n-1})$; furthermore, if $1 \leq m < n$, then

$$\begin{aligned}(\lambda_n I - T)\psi_m &= (\lambda_n - \lambda_m)\psi_m + (\lambda_m I - T)\psi_m \\ &\in \mathfrak{L}(\psi_1, \ldots, \psi_m) \subseteq \mathfrak{L}(\psi_1, \ldots, \psi_{n-1}).\end{aligned}$$

This, with (11), shows that

$$(\lambda_n I - T)\varphi_n \in \mathfrak{L}(\psi_1, \ldots, \psi_{n-1}) = \mathfrak{L}(\varphi_1, \ldots, \varphi_{n-1});$$

and so

$$0 = \langle (\lambda_n I - T)\varphi_n, \varphi_n \rangle = \lambda_n - \langle T\varphi_n, \varphi_n \rangle.$$

We have now constructed an orthonormal sequence (φ_n) such that $\langle T\varphi_n, \varphi_n \rangle = \lambda_n$. The closed subspace M generated by the φ_n's is the same as the one generated by the ψ_n's. Since each ψ_n is a principal vector of T associated with a non-zero eigenvalue, and each such principal vector is a linear combination of ψ_n's, it

follows that $M = \mathcal{P}_T$. Hence (φ_n) is an orthonormal basis of $\mathcal{P}_T$, and the theorem is proved.

THEOREM 1.8.7. *Suppose that* $\mathcal{H}$ *is a Hilbert space and* $T \in \mathcal{B}(\mathcal{H})$. *Then the following three conditions are equivalent.*

(i) T *is compact.*

(ii) Given any orthonormal system $\{\psi_k : k \in K\}$ *in* $\mathcal{H}$, $\langle T\psi_k, \psi_k \rangle \to 0$ *as* $k \to \infty$ *(in the sense of §1.2).*

(iii) There is a sequence (F_n) *of operators of finite rank on* $\mathcal{H}$, *such that* $||T-F_n|| \to 0$ *as* $n \to \infty$.

Proof. Suppose that (i) is satisfied and (ii) is not. For some orthonormal system $\{\psi_k : k \in K\}$ and some positive δ, the set $\{k \in K : |\langle T\psi_k, \psi_k \rangle| \geqslant 2\delta\}$ is infinite; it therefore has a countably infinite subset which, by change of notation, we can take to be the set of all positive integers. Thus $\{\psi_n : n = 1, 2, \ldots\}$ is an orthonormal sequence, and

$$(12) \qquad |\langle T\psi_n, \psi_n \rangle| \geqslant 2\delta \quad (n = 1, 2, \ldots).$$

Since T is compact, there is a subsequence $(n(j))$ of the positive integers such that $(T\psi_{n(j)})$ converges to an element x of $\mathcal{H}$. By deleting a finite number of terms of this sequence, we may suppose that

$$||T\psi_{n(j)} - x|| < \delta \quad (j = 1, 2, \ldots).$$

Thus

$$\begin{aligned} |\langle T\psi_{n(j)}, \psi_{n(j)} \rangle - \langle x, \psi_{n(j)} \rangle| &= |\langle T\psi_{n(j)} - x, \psi_n(j) \rangle| \\ &\leqslant ||T\psi_{n(j)} - x|| < \delta, \end{aligned}$$

and so, by (12), $|\langle x, \psi_{n(j)} \rangle| > \delta$ $(j = 1, 2, \ldots)$. Hence the series $\sum |\langle x, \psi_{n(j)} \rangle|^2$ diverges, contradicting Bessel's inequality. This shows that (i) implies (ii).

Suppose that (ii) is satisfied. Given a positive integer n, consider the class $\mathcal{S}$ of all orthonormal sets A in $\mathcal{H}$ for which

$$|\langle Ta, a \rangle| \geqslant \frac{1}{4n} \quad (a \in A)$$

(we allow the empty set as one possible choice of A). By (ii), each A in $\mathcal{S}$ is a finite set. Since the union of a strictly increasing sequence of sets in $\mathcal{S}$ is again a member of $\mathcal{S}$ (and therefore finite) each such sequence terminates. It follows that $\mathcal{S}$ has a maximal element B. If M is the (finite-dimensional) subspace of $\mathcal{H}$ generated by B, then $|\langle Tx, x \rangle| < (4n)^{-1}$ whenever $x \in M^\perp$ and $||x|| = 1$; for otherwise $\mathcal{S}$ contains $B \cup \{x\}$, contradicting the maximality of B. Thus $|\langle Tx, x \rangle| < n^{-1}$ whenever $x \in M^\perp$ and $||x|| \leqslant 2$; it now follows from equation 1.7(2) that

$$(13) \qquad |\langle Tu, v \rangle| \leqslant n^{-1} \quad (u, v \in M^\perp, ||u|| \leqslant 1, ||v|| \leqslant 1).$$

By taking $u = (I-P)x$ and $v = (I-P)y$, where P is the projection from $\mathcal{H}$ onto M, we deduce from (13) that

$$|\langle (I-P)T(I-P)x, y \rangle| \leqslant n^{-1}$$

whenever $x, y \in \mathcal{H}$, $||x|| \leqslant 1$ and $||y|| \leqslant 1$. Thus $||(I-P)T(I-P)|| \leqslant n^{-1}$. The operator

$$F_n = PT + TP - PTP$$

has finite rank, and $||T-F_n|| \leqslant n^{-1}$. This shows that (ii) implies (iii).

Since operators of finite rank are compact, and the set of compact linear operators is closed in $\mathcal{B}(\mathcal{H})$, it is apparent that (iii) implies (i). This completes the proof of the theorem.

It is not known whether or not conditions (i) and (iii) in Theorem 1.8.7 are equivalent for bounded linear operators acting on a general Banach space; for most of the particular Banach spaces encountered in elementary analysis, an affirmative answer is known.

THEOREM 1.8.8. *Suppose that T is a quasi-nilpotent operator acting on a Hilbert space $\mathcal{H}$, and the skew-adjoint part* Im T *is compact. Then T is compact.*

Proof. Let $A = \text{Re } T$, $B = \text{Im } T$, so that A and B are self-adjoint $T = A+iB$, and B is compact; we have to show that A is compact. Suppose the contrary. By Theorem 1.8.7, there is an orthonormal system $\{\psi_k : k \in K\}$ in $\mathcal{H}$ such that $\langle A\psi_k, \psi_k\rangle \not\to 0$ as $k \to \infty$. For some positive δ, the set

$$K_0 = \{k \in K : |\langle A\psi_k, \psi_k\rangle| \geqslant \delta\}$$

is infinite. If $\{\varphi_n : n = 1, 2, \ldots\}$ is a countably infinite subset of $\{\psi_k : k \in K_0\}$, then (φ_n) is an orthonormal sequence and

$$(14) \qquad |\langle A\varphi_n, \varphi_n\rangle| \geqslant \delta \quad (n = 1, 2, \ldots).$$

Since

$$|\langle A^m\varphi_n, \varphi_n\rangle|^2 \leqslant ||A^m\varphi_n||^2 = \langle A^m\varphi_n, A^m\varphi_n\rangle = \langle A^{2m}\varphi_n, \varphi_n\rangle$$

for all positive integers m and n, it follows from (14) that

$$(15) \qquad |\langle A^m\varphi_n, \varphi_n\rangle| \geqslant \delta^m \quad (n = 1, 2, \ldots)$$

whenever $m = 2^q$ for some $q = 0, 1, 2, \ldots$ For sufficiently large m of the form 2^q, $||T^m||^{1/m} < \delta$, and thus

$$\delta^m > ||T^m|| = ||(A+iB)^m|| = ||A^m - C||$$

for some compact linear operator C. By Theorem 1.8.7, $\langle C\varphi_n, \varphi_n\rangle \to 0$ as $n \to \infty$. Hence

$$|\langle A^m\varphi_n, \varphi_n\rangle| \leqslant |\langle (A^m - C)\varphi_n, \varphi_n\rangle| + |\langle C\varphi_n, \varphi_n\rangle| \leqslant ||A^m - C|| + |\langle C\varphi_n, \varphi_n\rangle| < \delta^m$$

for sufficiently large n. This contradicts (15), and so completes the proof of the theorem.

We conclude this section by describing a class of examples of compact linear operators. Suppose that $(E, \mathcal{E}, \mu)$ is a σ-finite measure space, $(E \times E, \mathcal{E} \times \mathcal{E}, \mu \times \mu)$ is the product of this measure space with itself, and $h \in L_2(E \times E, \mu \times \mu)$; for the terminology used here, we refer to [25: Chapter 7]. By Fubini's theorem,

$$(16) \qquad ||h||^2 = \int_E d\mu(s) \int_E |h(s, t)|^2 d\mu(t)$$

(here, and subsequently, the norm of any function refers to the usual norm in the appropriate L_2 space). For almost all s in E, $h(s, t)$ is of class $L_2(E, \mu)$ as a function of t; let Z denote the exceptional set of measure zero. If $f \in L_2(E, \mu)$, it follows from the Cauchy–Schwarz inequality that the integral

$$(T_h f)(s) = \int_E h(s, t) f(t) d\mu(t)$$

exists whenever $s \in E \sim Z$, and

$$(17) \qquad |(T_h f)(s)|^2 \leqslant ||f||^2 \int_E |h(s, t)|^2 d\mu(t).$$

It is readily verified that the function $T_h f$ (defined arbitrarily on Z) is measurable. From (16) and (17),

$$\int_E |(T_h f)(s)|^2 d\mu(s) \leqslant ||f||^2 \int d\mu(s) \int_E |h(s,\ t)|^2 d\mu(t) = ||f||^2\ ||h||^2,$$

so $T_h f \in L_2(E,\ \mu)$ and $||T_h f|| \leqslant ||h||\ ||f||$. It follows that T_h is a bounded linear operator on the Hilbert space $L_2(E,\ \mu)$, with $||T_h|| \leqslant ||h||$. We shall refer to h as an L_2 *kernel*, and to T_h as its associated *integral operator*. It follows easily from Fubini's theorem that, if $f,\ g \in L_2(E,\ \mu)$, then

$$\langle T_h f,\ g\rangle = \iint_{E\times E} h(s,\ t)\ \overline{g(s)}\ f(t) d\mu(s) d\mu(t). \tag{18}$$

From this, it is readily verified that the adjoint T_h^* is the integral operator associated with the L_2 kernel h^*, where $h^*(s,\ t) = \overline{h(t,\ s)}$.

We assert that T_h is a compact linear operator. To prove this, suppose that $\{\psi_k : k \in K\}$ is an orthonormal system in $L_2(E,\ \mu)$. The functions Ψ_k, defined on $E \times E$ by $\Psi_k(s,\ t) = \psi_k(s)\overline{\psi_k(t)}$, form an orthonormal system in $L_2(E \times E,\ \mu \times \mu)$. With $f = g = \psi_k$, it follows from (18) that

$$\langle T_h \psi_k,\ \psi_k\rangle = \langle h,\ \Psi_k\rangle,$$

and Bessel's inequality asserts that

$$\sum_{k\in K} |\langle T_h\psi_k,\ \psi_k\rangle|^2 = \sum_{k\in K} |\langle h,\ \Psi_k\rangle|^2 \leqslant ||h||^2.$$

Thus $\langle T_h\psi_k,\ \psi_k\rangle \to 0$ as $k \to \infty$. By Theorem 1.8.7, T_h is a compact linear operator.

Finally, we assert that $T_h = 0$ if and only if $h(s,\ t) = 0$ almost everywhere on $E \times E$. The 'if' part of this statement is an

immediate consequence of the inequality $||T_h|| \leqslant ||h||$. Now suppose that $T_h = 0$. To prove that $h(s,\ t) = 0$ almost everywhere on $E \times E$, it is sufficient to show that $h(s,\ t) = 0$ almost everywhere on $E_0 \times E_0$, whenever E_0 is a measurable subset of E with $\mu(E_0) < \infty$. Let $\mathcal{S}$ denote the class of all measurable subsets S of $E_0 \times E_0$ which satisfy

$$\iint_S h(s,\ t) d\mu(s) d\mu(t) = 0. \tag{19}$$

As a function on $E_0 \times E_0$, h is of class L_2, and therefore (since $E_0 \times E_0$ has finite measure) of class L_1. If $S_1,\ S_2,\ S_3,\ \ldots \in \mathcal{S}$, and the sequence (S_n) is either increasing or decreasing, it follows easily from the dominated convergence theorem that $\lim S_n \in \mathcal{S}$. Following [25: p. 27] we describe this property by saying that $\mathcal{S}$ is a *monotone* class. If A and B are measurable subsets of E_0, we can take f and g in (18) to be the characteristic functions of B and A respectively, and deduce that $A \times B \in \mathcal{S}$. It follows that $\mathcal{S}$ contains the ring consisting of all finite disjoint unions of such sets $A \times B$. Since $\mathcal{S}$ is monotone, it contains the σ-ring generated by this last ring [25: p. 27]; that is, $\mathcal{S}$ consists of *all* measurable subsets S of $E_0 \times E_0$, and (19) is satisfied for all such S. By taking for S, in turn, the four sets on which the real and imaginary parts of $h(s,\ t)$ both have constant sign, it follows from (19) that $h(s,\ t) = 0$ almost everywhere on $E_0 \times E_0$.

1.9. *Compact normal operators*

We begin this section by describing a simple process for constructing compact linear operators, in particular, compact normal operators. Suppose that $\mathcal{H}$ is a Hilbert space, $(\varphi_1,\ \varphi_2,\ \ldots)$ and $(\psi_1,\ \psi_2,\ \ldots)$ are othonormal sequences in $\mathcal{H}$, and $(\lambda_1,\ \lambda_2,\ \ldots)$ is a sequence of non-zero complex numbers satisfying

(1) $$|\lambda_1| \geqslant |\lambda_2| \geqslant |\lambda_3| \geqslant \dots$$

We require that the sequences (λ_j), (φ_j), (ψ_j) all have the same number of terms (finite or countably infinite); when they are infinite, we suppose that (λ_n) converges to 0.

For each x in $\mathcal{H}$,

$$\Sigma\, |\lambda_j \langle x, \varphi_j \rangle|^2 \leqslant |\lambda_1|^2 \Sigma\, |\langle x, \varphi_j \rangle|^2 \leqslant |\lambda_1|^2 ||x||^2.$$

Hence the equation

(2) $$Tx = \Sigma\, \lambda_j \langle x, \varphi_j \rangle \psi_j$$

defines an element Tx of $\mathcal{H}$, and $||Tx|| \leqslant |\lambda_1|\; ||x||$. It follows that T is a bounded linear operator on $\mathcal{H}$, and $||T|| \leqslant |\lambda_1|$. Furthermore

$$T\varphi_j = \lambda_j \psi_j \quad (j = 1, 2, \dots);$$

by taking $j = 1$, we deduce that $||T|| \geqslant |\lambda_1|$, so

(3) $$||T|| = |\lambda_1|.$$

We assert that T is a compact linear operator. If the sequence (λ_j) is finite, then T has finite rank and the result is apparent. If (λ_j) is an infinite sequence then, for each $k = 1, 2, \dots$, we can define an operator T_k of finite rank by

$$T_k x = \sum_{j=1}^{k} \lambda_j \langle x, \varphi_j \rangle \psi_j \quad (x \in \mathcal{H}).$$

Since

$$(T-T_k)x = \sum_{j=k+1}^{\infty} \lambda_j \langle x, \varphi_j \rangle \psi_j \quad (x \in \mathcal{H})$$

it follows from (3), with $T-T_k$ in place of T, that

$$||T-T_k|| = |\lambda_{k+1}| \to 0$$

as $k \to \infty$. By Theorem 1.8.7, T is compact.

The conditions imposed in the first paragraph of §1.9 remain satisfied when (φ_j) and (ψ_j) are interchanged and (λ_j) is replaced by the complex conjugate sequence $(\bar{\lambda}_j)$. It follows that the equation

(4) $$T'y = \Sigma\, \bar{\lambda}_j \langle y, \psi_j \rangle \varphi_j \quad (y \in \mathcal{H})$$

defines a bounded linear operator T' on $\mathcal{H}$. By (4) and (2),

$$\langle T'y, x \rangle = \Sigma\, \bar{\lambda}_j \langle y, \psi_j \rangle \langle \varphi_j, x \rangle,$$

$$\begin{aligned}\langle x, T^*y \rangle &= \langle Tx, y \rangle \\ &= \Sigma\, \lambda_j \langle x, \varphi_j \rangle \langle \psi_j, y \rangle \\ &= \overline{\langle T'y, x \rangle} \\ &= \langle x, T'y \rangle \quad (x, y \in \mathcal{H}).\end{aligned}$$

Thus $T^* = T'$, and therefore

(5) $$T^*y = \Sigma\, \bar{\lambda}_j \langle y, \psi_j \rangle \varphi_j \quad (y \in \mathcal{H}).$$

Now suppose that $\varphi_j = \psi_j$ $(j = 1, 2, \dots)$. It then follows from (5) and (2) that

$$||T^*x||^2 = \Sigma\, |\lambda_j \langle x, \varphi_j \rangle|^2 = ||Tx||^2 \quad (x \in \mathcal{H});$$

and, by Lemma 1.7.3, T is normal.

We have now shown that, if the sequences (φ_j), (ψ_j) and (λ_j) satisfy the conditions set out in the first paragraph of this section, then equation (2) defines a compact linear operator T; and T is normal if $\varphi_j = \psi_j$ for all j. Our main purpose in §1.9 is to show

that every compact normal operator arises in this way. From this, we shall deduce that each compact linear operator can be represented in the form (2), by an appropriate choice of λ_j, φ_j and ψ_j. Before proving these results, we require the following lemma.

LEMMA 1.9.1. *Let T be a compact normal operator acting on a Hilbert space $\mathcal{H}$.*

(i) T has an eigenvalue λ_1 satisfying $|\lambda_1| = ||T||$.

(ii) Every eigenvalue of T has index 1.

(iii) If λ is a scalar and $x \in \mathcal{H}$, then $Tx = \lambda x$ if and only if $T^*x = \bar{\lambda}x$.

(iv) If x and y are eigenvectors of T, associated with distinct eigenvalues λ and μ respectively, then $\langle x, y\rangle = 0$.

Proof. By Theorem 1.7.4, the spectral radius of T is $||T||$, so there is an element λ_1 of $\sigma(T)$ satisfying $|\lambda_1| = ||T||$. If $T \neq 0$, then $\lambda_1 \neq 0$, and λ_1 is an eigenvalue of T by Theorem 1.8.2. This proves (i), since the stated result is obvious when $T = 0$.

By Lemma 1.7.3, $||Tx|| = ||T^*x||$ for each x in $\mathcal{H}$. With T replaced by the normal operator $T-\lambda I$, we have

$$||Tx-\lambda x|| = ||T^*x-\bar{\lambda}x||,$$

which proves (iii).

Suppose that λ is an eigenvalue of T, $x \in \mathcal{H}$, and $(T-\lambda I)^2x = 0$. By (iii), with $Tx-\lambda x$ in place of x, we have

$$(T^*-\bar{\lambda}I)(T-\lambda I)x = 0.$$

Thus

$$||(T-\lambda I)x||^2 = \langle (T^*-\bar{\lambda}I)(T-\lambda I)x,\ x\rangle = 0,$$

and therefore $(T-\lambda I)x = 0$. This proves (ii).

Under the conditions set out in (iv),

$$\begin{aligned}\lambda\langle y, x\rangle = \langle y, \bar{\lambda}x\rangle &= \langle y, T^*x\rangle\\ &= \langle Ty, x\rangle\\ &= \langle \mu y, x\rangle = \mu\langle y, x\rangle.\end{aligned}$$

Since $\lambda \neq \mu$, it follows that $\langle y, x\rangle = 0$.

Suppose that T is a compact normal operator acting on a Hilbert space $\mathcal{H}$. By Theorem 1.8.2, the distinct non-zero eigenvalues of T can be arranged as a (finite or infinite) sequence $(\mu_1, \mu_2, \ldots)$; and, in the infinite case, (μ_n) converges to 0. We may assume that the μ_j's have been ordered so that

$$(6)\qquad |\mu_1| \geqslant |\mu_2| \geqslant |\mu_3| \geqslant \ldots$$

The algebraic multiplicity $m(j)$ of μ_j as an eigenvalue of T coincides with the geometric multiplicity, by Lemma 1.9.1(ii). Let (λ_n) be the sequence consisting of $m(1)$ terms μ_1, then $m(2)$ terms μ_2, then $m(3)$ terms μ_3, and so on. If (μ_n) is a finite sequence, so is (λ_n); in the infinite case, (λ_n) converges to 0.

The null space $\mathfrak{N}(j)$ of μ_jI-T has dimension $m(j)$ and so has an orthonormal basis $e(j, 1)$, $e(j, 2)$, $\ldots$, $e(j, m(j))$. Let (φ_n) be the sequence of vectors obtained by taking first $e(1, 1), \ldots, e(1, m(1))$, then $e(2, 1), \ldots, e(2, m(2))$, then $e(3, 1), \ldots, e(3, m(3))$, and so on. For any given j, the integers n such that $\lambda_n = \mu_j$ are precisely those for which φ_n has the form $e(j, \ell)$ $(1 \leqslant \ell \leqslant m(j))$; so

$$T\varphi_n = \lambda_n\varphi_n.$$

If $j \neq k$ then $\langle e(j, \ell), e(k, m)\rangle = 0$, by Lemma 1.9.1(iv); since $e(j, 1), \ldots, e(j, m(j))$ is orthonormal, for each j, it follows that (φ_n) is an orthonormal sequence. We shall describe (λ_n) as the sequence

of non-zero eigenvalues of T, *arranged in order of decreasing magnitude and counted according to their multiplicities*; (φ_n) is the *corresponding orthonormal sequence of eigenvectors.* In general, the sequence (λ_n) is not uniquely determined; for distinct eigenvalues of T may have the same absolute value, and then (6) does not completely specify the order of the μ_j's. However, (λ_n) is uniquely determined if all the μ_j's are positive; we shall see later that this last condition is equivalent to the statement that T is a positive operator.

If we take $\psi_n = \varphi_n$, the conditions set out in the first paragraph of §1.9 are satisfied. It follows that the equation

(7) $$T_0 x = \Sigma\, \lambda_n \langle x, \varphi_n\rangle \varphi_n \quad (x \in \mathcal{H})$$

defines a compact normal operator T_0 on $\mathcal{H}$. We shall show that $T = T_0$.

Let M be the closed subspace generated by the vectors φ_n, and denote by P the projection from $\mathcal{H}$ onto M. For each n,

$$T_0\varphi_n = \lambda_n\varphi_n = T\varphi_n;$$

by linearity and continuity, $T_0 y = Ty$ whenever $y \in M$. Thus $T_0P = TP$. However,

$$\begin{aligned} T_0 x &= \Sigma\, \lambda_n \langle x, \varphi_n\rangle\varphi_n \\ &= \Sigma\, \lambda_n\langle x, P\varphi_n\rangle\varphi_n \\ &= \Sigma\, \lambda_n\langle Px, \varphi_n\rangle\varphi_n \\ &= T_0Px \quad (x \in \mathcal{H}), \end{aligned}$$

so $T_0 = T_0P$. Thus

(8) $$TP = T_0.$$

Also, since $PTx \in M$ whenever $x \in \mathcal{H}$, and (φ_n) is an orthonormal basis of M,

$$\begin{aligned} PTx &= \Sigma \langle PTx, \varphi_n\rangle\varphi_n \\ &= \Sigma\langle x, T^*P\varphi_n\rangle\varphi_n. \end{aligned}$$

Since $T^*P\varphi_n = T^*\varphi_n = \bar{\lambda}_n\varphi_n$, by Lemma 1.9.1(iii), it follows that

$$PTx = \Sigma\, \lambda_n\langle x, \varphi_n\rangle\varphi_n = T_0x.$$

This, with (8), shows that $TP = T_0 = PT$. By taking adjoints, we obtain $PT^* = T^*P$. Since T is normal, it follows that the operators P, T, T^* all commute. Hence $T-T_0$ ($= T-TP$) is normal.

Suppose that $T \neq T_0$. By Lemma 1.9.1(i), $T-T_0$ has a non-zero eigenvalue μ. Let φ ($\neq 0$) be a corresponding eigenvector. Since $(T-T_0)\varphi_n = 0$ and $(T-T_0)\varphi = \mu\varphi$, it follows from Lemma 1.9.1(iv) that

(9) $$\langle\varphi, \varphi_n\rangle = 0 \quad (n = 1, 2, \ldots).$$

By (7), $T_0\varphi = 0$, so $T\varphi = (T-T_0)\varphi = \mu\varphi$. Thus μ is a non-zero eigenvalue of T, so $\mu = \mu_j$ for some j, and $\varphi \in \mathfrak{N}(j)$. However, $\mathfrak{N}(j)$ has a basis consisting of a finite set of φ_n's, and by (9), $\varphi \in \mathfrak{N}(j)^{\perp}$. Hence $\varphi \in \mathfrak{N}(j) \cap \mathfrak{N}(j)^{\perp} = \{0\}$, contradicting our assumption that $\varphi \neq 0$. It follows that $T = T_0$.

The results of the preceding paragraphs are summarized in the following theorem.

THEOREM 1.9.2. *Suppose that T is a compact normal operator acting on a Hilbert space $\mathcal{H}$, (λ_n) is the sequence of non-zero eigenvalues (arranged in order of decreasing magnitude and counted according to their multiplicities) and (φ_n) is the corresponding orthonormal sequence of eigenvectors. Then*

$$Tx = \Sigma\, \lambda_n\langle x, \varphi_n\rangle\varphi_n \quad (x \in \mathcal{H}).$$

We have shown that every compact normal operator T can be represented in the form (2), with $\psi_j = \varphi_j$ and (λ_j) the sequence of non-zero eigenvalues of T. This applies, in particular, to compact self-adjoint operators. By comparing (2) and (5), in the case where $\varphi_j = \psi_j$, it follows that a compact normal operator T is self-adjoint if and only if each of its eigenvalues is real. By (2), with $\varphi_j = \psi_j$,

$$\langle Tx,\ x\rangle = \Sigma\ \lambda_j |\langle x,\ \varphi_j\rangle|^2 \quad (x \in \mathcal{H}),$$

so T is positive if and only if the λ_n's are all positive.

The functional calculus for a compact self-adjoint operator (see Theorem 1.7.7 and the remarks following its proof) assumes a very simple form. We may assume that T is represented as in Theorem 1.9.2, with all the λ_n's real. Suppose that $f \in C(\sigma(T))$. If $\langle z,\ \varphi_n\rangle = 0$ for each n, then $Tz = 0$, $T\varphi_n = \lambda_n\varphi_n$; it follows from Theorem 1.7.7(vii) that $f(T)z = f(0)z$ and $f(T)\varphi_n = f(\lambda_n)\varphi_n$. Every vector x has the form

$$x = \Sigma\ \langle x,\ \varphi_n\rangle\varphi_n + z,$$

where $\langle z,\ \varphi_n\rangle = 0$ for all n. Thus

$$\begin{aligned} f(T)x &= \Sigma\ \langle x,\ \varphi_n\rangle f(T)\varphi_n + f(T)z \\ &= \Sigma\ f(\lambda_n)\ \langle x,\ \varphi_n\rangle\varphi_n + f(0)z. \end{aligned}$$

The case in which $f(0) = 0$ is of particular interest. In this case,

$$f(T)x = \Sigma\ f(\lambda_n)\ \langle x,\ \varphi_n\rangle\varphi_n \quad (x \in \mathcal{H}). \tag{10}$$

If the sequence (λ_n) is infinite, it converges to 0; by continuity, $f(\lambda_n) \to f(0) = 0$. If follows that $f(T)$ is a compact normal operator if $f(0) = 0$.

We now revert to the study of a general (not necessarily normal) compact linear operator T. By use of Theorem 1.9.2, together with the polar decomposition of T, we shall show that T can be expressed in the form (2), with sequences (λ_n), (φ_n), (ψ_n) satisfying the conditions set out in the first paragraph of §1.9.

Suppose that $\mathcal{H}$ is a Hilbert space and, for each T in $\mathcal{B}(\mathcal{H})$, let $\mathcal{R}_T$ denote the closed range space of T. We recall from Theorem 1.7.12 that there exist a positive linear operator H $(= (T^*T)^{1/2})$, and a partial isometry V with initial space $\mathcal{R}_{T^*}$ $(= \mathcal{R}_H)$ and final space $\mathcal{R}_T$, such that $T = VH$, $H = V^*T$. From these relations it follows that T is compact if and only if H is compact; and T has finite rank if and only if H has (the same) finite rank.

THEOREM 1.9.3. *Suppose that (μ_n) is a decreasing sequence of positive real numbers which is either finite or infinite and convergent to 0, and that (φ_n) and (ψ_n) are orthonormal sequences in a Hilbert space $\mathcal{H}$. Then the equation*

$$Tx = \Sigma\ \mu_n\langle x,\ \varphi_n\rangle\psi_n \quad (x \in \mathcal{H}) \tag{11}$$

defines a compact linear operator T on $\mathcal{H}$; furthermore, T has finite rank k if and only if the sequence (μ_n) terminates after just k terms.

Conversely, if T is a compact linear operator acting on $\mathcal{H}$, then T can be expressed in the above form; the sequence (μ_n) is uniquely determined, and consists of the eigenvalues of $(T^*T)^{1/2}$, arranged in order of decreasing magnitude and counted according to their multiplicities.

Proof. We have already seen, at the beginning of §1.9, that equation (11) defines a compact linear operator T, with adjoint T^* satisfying

$$T^*y = \Sigma\ \mu_n\langle y,\ \psi_n\rangle\varphi_n \quad (y \in \mathcal{H}). \tag{12}$$

Thus $T^*\psi_n = \mu_n\varphi_n$ and, by (11),

$$T^*Tx = \Sigma \mu_n\langle x, \varphi_n\rangle T^*\psi_n = \Sigma\, \mu_n^2\langle x, \varphi_n\rangle\varphi_n \quad (x \in \mathcal{H}).$$

With f defined on the spectrum of the positive operator T^*T by $f(t) = t^{1/2}$, it follows from (10) (with T^*T and μ_n^2 in place of T and λ_n, respectively) that the operator $H = (T^*T)^{1/2}$ satisfies

$$Hx = \Sigma\, \mu_n\langle x, \varphi_n\rangle\varphi_n \quad (x \in \mathcal{H}). \tag{13}$$

From (13), and the assumption that (μ_n) decreases, it is easily verified that (μ_n) is the sequence of non-zero eigenvalues of H, in decreasing order and counted according to their multiplicities. Since H is a positive operator, this implies that (μ_n) is uniquely determined when T (hence H) is given. It is clear from (11) that (ψ_n) is an orthonormal basis of $\mathcal{R}_T$; so T has finite rank k if and only if (μ_n) terminates after just k terms.

Conversely, let T be a compact linear operator acting on $\mathcal{H}$, and let its polar decomposition be VH, so that H $(=(T^*T)^{1/2})$ is positive and compact. Let (μ_n) be the sequence of non-zero eigenvalues of H (in decreasing order, and counted according to their multiplicities). By Theorem 1.9.2, there is an orthonormal sequence (φ_n) such that (13) is satisfied. Thus

$$Tx = VHx = \Sigma\, \mu_n\langle x, \varphi_n\rangle\psi_n \quad (x \in \mathcal{H})$$

where $\psi_n = V\varphi_n$. Now V is a partial isometry with initial space $\mathcal{R}_H$, and φ_n $(= \mu_n^{-1}H\varphi_n)$ lies in $\mathcal{R}_H$; so $V^*V\varphi_n = \varphi_n$. Thus

$$\langle\psi_m, \psi_n\rangle = \langle V\varphi_m, V\varphi_n\rangle = \langle V^*V\varphi_m, \varphi_n\rangle = \langle\varphi_m, \varphi_n\rangle.$$

Hence (ψ_n) is an orthonormal sequence, and the theorem is proved.

If $\mathcal{H}$ is a Hilbert space and $x, y \in \mathcal{H}$, we define a bounded linear operator $x \otimes y$ on $\mathcal{H}$ by $(x \otimes y)z = \langle z, y\rangle x$ $(z \in \mathcal{H})$. It is clear that $\|x \otimes y\| = \|x\|\,\|y\|$, and that $x \otimes y$ has rank 1 if $x \neq 0 \neq y$. Now suppose that $T \in \mathcal{B}(\mathcal{H})$ and T has finite rank m. It follows from Theorem 1.9.3 that T can be expressed in the form

$$T = \sum_{n=1}^{m} \mu_n\psi_n \otimes \varphi_n,$$

where (λ_n) is a sequence of positive real numbers (the eigenvalues of $(T^*T)^{1/2}$) and (φ_n), (ψ_n) are orthonormal sequences. By (12),

$$T^* = \sum_{n=1}^{m} \mu_n\varphi_n \otimes \psi_n,$$

and so T^* has the same finite rank as T. If T has polar decomposition VH then, by (13),

$$H = \sum_{n=1}^{m} \mu_n\varphi_n \otimes \varphi_n.$$

CHAPTER 2

The von Neumann-Schatten Classes of Operators

2.1. *Introduction*

This chapter is concerned with certain classes $\mathcal{C}_p$ $(1 \leqslant p < \infty)$ of linear operators on a Hilbert space $\mathcal{H}$ which were introduced by von Neumann and Schatten [49]. It turns out that each of these classes is a two sided ideal in $\mathcal{B}(\mathcal{H})$, and consists of compact operators. When provided with a suitable norm, $\mathcal{C}_p$ becomes a Banach space with properties closely analagous to those of the sequence space ℓ_p (see §1.3). The development of this theory involves difficulties, arising from the non-commutativity of $\mathcal{C}_p$, which have no parallel in the case of ℓ_p.

Throughout Chapter 2, $\mathcal{H}$ denotes a Hilbert space and J is a set whose cardinality is the dimension of $\mathcal{H}$. Hence each orthonormal basis of $\mathcal{H}$ can be expressed as $\{\varphi_j : j \in J\}$, while a general (not necessarily complete) orthonormal system has the form $\{\psi_k : k \in K\}$, where the cardinality of K is not greater than that of J.

DEFINITION 2.1.1. When $1 \leqslant p < \infty$, $\mathcal{C}_p$ is the set of all operators T in $\mathcal{B}(\mathcal{H})$ which satisfy the following condition: for each orthonormal system $\{\psi_k : k \in K\}$ in $\mathcal{H}$,

$$\sum_{k \in K} |\langle T\psi_k, \psi_k \rangle|^p < \infty .$$

We shall adopt the convention that $\mathcal{C}_\infty$ is $\mathcal{B}(\mathcal{H})$. It is apparent from the properties of ℓ_p spaces which are described in §1.3 that $\mathcal{C}_p$ is a linear subspace of $\mathcal{B}(\mathcal{H})$, and that $\mathcal{C}_p \subseteq \mathcal{C}_q$ if $1 \leqslant p \leqslant q \leqslant \infty$. Furthermore $T^* \in \mathcal{C}_p$ whenever $T \in \mathcal{C}_p$, and so $\mathcal{C}_p$ contains the self-adjoint and skew-adjoint parts of each of its members. It follows from Theorem 1.8.7 that, if $1 \leqslant p < \infty$, then each element of $\mathcal{C}_p$ is a compact operator.

LEMMA 2.1.2. *Suppose that $1 \leqslant p < \infty$, T is a compact self-adjoint operator on $\mathcal{H}$, and (λ_n) is the sequence of non-zero eigenvalues of T, counted according to their multiplicities.*

(i) *If $T \in \mathcal{C}_p$ then $\sum_n |\lambda_n|^p < \infty$.*

(ii) *If $\sum_n |\lambda_n|^p < \infty$, then $T \in \mathcal{C}_p$ and, for each orthonormal system $\{\psi_k : k \in K\}$ in $\mathcal{H}$,*

(1) $$\sum_{k \in K} |\langle T\psi_k, \psi_k\rangle|^p \leqslant \sum_n |\lambda_n|^p.$$

Proof. There is an orthonormal sequence (φ_n) in $\mathcal{H}$ for which

$$Tx = \sum_n \lambda_n \langle x, \varphi_n\rangle \varphi_n \quad (x \in \mathcal{H}).$$

If $T \in \mathcal{C}_p$ then, by Definition 2.1.1,

$$\sum_n |\lambda_n|^p = \sum_n |\langle T\varphi_n, \varphi_n\rangle|^p < \infty.$$

Suppose conversely that $\sum |\lambda_n|^p < \infty$. If $\{\psi_k : k \in K\}$ is any orthonormal system in $\mathcal{H}$, then

$$\langle T\psi_k, \psi_k\rangle = \sum_n \lambda_n |\langle \psi_k, \varphi_n\rangle|^2.$$

With q the index conjugate to p, Hölder's inequality (with the usual interpretation when $p = 1$ and $q = \infty$) yields

$$\begin{aligned}|\langle T\psi_k, \psi_k\rangle| &\leqslant \sum_n |\lambda_n| \, |\langle \psi_k, \varphi_n\rangle|^{2/p} |\langle \psi_k, \varphi_n\rangle|^{2/q} \\ &\leqslant [\sum_n |\lambda_n|^p |\langle \psi_k, \varphi_n\rangle|^2]^{1/p} [\sum_n |\langle \psi_k, \varphi_n\rangle|^2]^{1/q}.\end{aligned}$$

Since $\sum_n |\langle \psi_k, \varphi_n\rangle|^2 \leqslant ||\psi_k||^2 = 1$, we have

$$\begin{aligned}\sum_k |\langle T\psi_k, \psi_k\rangle|^p &\leqslant \sum_k \sum_n |\lambda_n|^p \, |\langle \psi_k, \varphi_n\rangle|^2 \\ &\leqslant \sum_n |\lambda_n|^p \sum_k |\langle \varphi_n, \psi_k\rangle|^2 \\ &\leqslant \sum_n |\lambda_n|^p \, ||\varphi_n||^2 = \sum_n |\lambda_n|^p.\end{aligned}$$

Thus $T \in \mathcal{C}_p$ and (1) is satisfied.

We have seen in Theorem 1.8.7 that an operator T on $\mathcal{H}$ is compact if and only if it is the limit in norm of operators of finite rank. It follows at once that T is compact if and only if there is a sequence (F_n) of operators on $\mathcal{H}$, for which F_n has finite rank not more than n and $\lim_{n\to\infty} ||T - F_n|| = 0$. Our next result shows that membership of $\mathcal{C}_p$ is equivalent to a qualitative statement about the rapidity with which $||T - F_n||$ can converge to 0.

THEOREM 2.1.3. *Suppose $T \in \mathcal{B}(\mathcal{H})$ and $1 \leqslant p < \infty$. Then $T \in \mathcal{C}_p$ if and only if there is a sequence (F_n) of operators on $\mathcal{H}$, such that F_n has finite rank not greater than n and*

$$\sum_{n=1}^{\infty} ||T - F_n||^p < \infty.$$

Proof. Let $\mathfrak{D}$ denote the class of all operators T on $\mathcal{H}$ for which such a sequence (F_n) exists. Given S and T in $\mathfrak{D}$ and scalars α and β, choose operators F_n and G_n with finite rank not more than n for which

$$(2) \qquad \sum_{n=1}^{\infty} ||S-F_n||^p < \infty, \qquad \sum_{n=1}^{\infty} ||T-G_n||^p < \infty,$$

and define a sequence (H_n) by

$$H_1 = 0, \qquad H_{2n} = H_{2n+1} = \alpha F_n + \beta G_n \quad (n = 1, 2, \ldots).$$

Then H_n has finite rank not greater than n, and

$$\begin{aligned}\sum_{n=2}^{\infty} ||\alpha S+\beta T-H_n||^p \\ &= 2\sum_{n=1}^{\infty} ||\alpha(S-F_n)+\beta(T-G_n)||^p \\ &< 2\sum_{n=1}^{\infty} \{|\alpha|\ ||S-F_n|| + |\beta|\ ||T-G_n||\}^p.\end{aligned}$$

The last quantity is finite, by (2) and Minkowski's inequality, so $\alpha S+\beta T \in \mathfrak{D}$. Furthermore $S^* \in \mathfrak{D}$, since F_n^* has finite rank not more than n and

$$\sum_{n=1}^{\infty} ||S^*-F_n^*||^p = \sum_{n=1}^{\infty} ||S-F_n||^p < \infty.$$

Finally, S is compact since it is the limit in norm of the sequence (F_n) of operators of finite rank.

It follows from the results just proved that $\mathfrak{D}$ is a linear subspace of $\mathcal{B}(\mathcal{H})$ which consists of compact operators and contains the self-adjoint and skew-adjoint parts of each of its members. We have already seen that $\mathcal{C}_p$ has these properties. Hence each of the subspaces $\mathfrak{D}$ and $\mathcal{C}_p$ is the linear span of its self-adjoint elements; and in order to prove that $\mathfrak{D} = \mathcal{C}_p$, it is sufficient to show that a compact self-adjoint operator T lies in $\mathfrak{D}$ if and only if it lies in $\mathcal{C}_p$. Given such a T, we may suppose that

$$Tx = \sum_{m=1}^{\infty} \lambda_m \langle x, \varphi_m \rangle \varphi_m \quad (x \in \mathcal{H}),$$

where (φ_m) is an orthonormal sequence and (λ_m) is the sequence of non-zero eigenvalues of T (arranged in order of decreasing magnitude, counted according to their algebraic multiplicities, and followed by an infinite sequence of zeros if T is of finite rank).

If $T \in \mathcal{C}_p$, define an operator F_p by

$$F_n x = \sum_{m=1}^{n} \lambda_m \langle x, \varphi_m \rangle \varphi_m \quad (x \in \mathcal{H}).$$

Then F_n has finite rank not greater than n,

$$(T-F_n)x = \sum_{m=n+1}^{\infty} \lambda_m \langle x, \varphi_m \rangle \varphi_m \quad (x \in \mathcal{H}),$$

$$||T-F_n|| = \sup\{|\lambda_m| : m > n\} = |\lambda_{n+1}|.$$

By Lemma 2.1.2,

$$\sum_{n=1}^{\infty} ||T-F_n||^p = \sum_{n=1}^{\infty} |\lambda_{n+1}|^p < \infty,$$

so $T \in \mathfrak{D}$.

Suppose conversely that $T \in \mathfrak{D}$. Let G_n be an operator of finite rank not more than n ($n = 1, 2, \ldots$) for which $\Sigma\, ||T-G_n||^p < \infty$. For a fixed n, we may choose vectors x_m, y_m ($1 \leqslant m \leqslant n$) in $\mathcal{H}$ such that

$$G_n x = \sum_{m=1}^{n} \langle x, x_m \rangle y_m \quad (x \in \mathcal{H}).$$

Then $G_n(\alpha_1\varphi_1 + \ldots + \alpha_{n+1}\varphi_{n+1}) = 0$ provided that the scalars $\alpha_1, \ldots, \alpha_{n+1}$ satisfy the n linear equations

$\langle \alpha_1\varphi_1+\ldots+\alpha_{n+1}\varphi_{n+1}, x_m\rangle = 0$ $(m = 1, \ldots, n)$. With $\alpha_1, \ldots, \alpha_{n+1}$ (not all zero) satisfying these equations, we have

$$\begin{aligned}||(T-G_n)(\alpha_1\varphi_1+\ldots+\alpha_{n+1}\varphi_{n+1})|| &= ||T(\alpha_1\varphi_1+\ldots+\alpha_{n+1}\varphi_{n+1})|| \\ &= ||\lambda_1\alpha_1\varphi_1+\ldots+\lambda_{n+1}\alpha_{n+1}\varphi_{n+1}|| \\ &= [|\lambda_1\alpha_1|^2+\ldots+|\lambda_{n+1}\alpha_{n+1}|^2]^{1/2} \\ &\geqslant |\lambda_{n+1}|[|\alpha_1|^2+\ldots+|\alpha_{n+1}|^2] \\ &= |\lambda_{n+1}|\ ||\alpha_1\varphi_1+\ldots+\alpha_{n+1}\varphi_{n+1}||.\end{aligned}$$

Hence $|\lambda_{n+1}| \leqslant ||T-G_n||$,

$$\sum_{n=1}^{\infty} |\lambda_n|^p \leqslant |\lambda_1|^p + \sum_{n=1}^{\infty} ||T-G_n||^p < \infty,$$

and by Lemma 2.1.2, $T \in \mathcal{C}_p$.

COROLLARY 2.1.4. *$\mathcal{C}_p$ contains each operator of finite rank on $\mathcal{H}$ $(1 \leqslant p \leqslant \infty)$.*

Proof. This is obvious when $p = \infty$, and follows at once from Theorem 2.1.3 when $1 \leqslant p < \infty$.

LEMMA 2.1.5. *Suppose $1 \leqslant p \leqslant \infty$, $T \in \mathcal{C}_p$ and $A \in \mathcal{B}(\mathcal{H})$. Then $AT \in \mathcal{C}_p$, $TA \in \mathcal{C}_p$.*

Proof. We may suppose $1 \leqslant p < \infty$, since the result is evident when $p = \infty$. By Theorem 2.1.3 there is a sequence (F_n) of operators on $\mathcal{H}$ for which F_n has finite rank not greater than n and $\Sigma\, ||T-F_n||^p < \infty$. Since AF_n and F_nA have finite rank not more than n and

$$\Sigma\, ||AT-AF_n||^p \leqslant ||A||^p\, \Sigma\, ||T-F_n||^p < \infty,$$
$$\Sigma\, ||TA-F_nA||^p \leqslant ||A||^p\, \Sigma\, ||T-F_n||^p < \infty,$$

it follows, again from Theorem 2.1.3, that $AT \in \mathcal{C}_p$ and $TA \in \mathcal{C}_p$.

Some of the results proved above are of purely technical interest, but will be needed in §2.3 when $\mathcal{C}_p$ is studied as a Banach space. We conclude this section by summarising, as a theorem, the main algebraic properties of $\mathcal{C}_p$ which have been established so far.

THEOREM 2.1.6. *Suppose $1 \leqslant p \leqslant \infty$. Then $\mathcal{C}_p$ is a two sided ideal in $\mathcal{B}(\mathcal{H})$. It contains each operator of finite rank on $\mathcal{H}$, and also contains the adjoint of each of its members. If $1 \leqslant p < \infty$, then each element of $\mathcal{C}_p$ is a compact operator.*

2.2. *The trace on $\mathcal{C}_1$*

If $\mathcal{H}$ is an n-dimensional Hilbert space and $\varphi_1, \ldots, \varphi_n$ is an orthonormal basis in $\mathcal{H}$, we can define a linear functional τ on $\mathcal{B}(\mathcal{H})$ by

$$\tau(A) = \langle A\varphi_1, \varphi_1\rangle + \ldots + \langle A\varphi_n, \varphi_n\rangle \quad (A \in \mathcal{B}(\mathcal{H})).$$

Elementary algebraic manipulations show that τ does not depend on the choice of the orthonormal basis $\varphi_1, \ldots, \varphi_n$; and that, if A has matrix $[a_{ij}]$ with respect to some orthonormal basis, then $\tau(A) = a_{11}+\ldots+a_{nn}$, the trace of the matrix $[a_{ij}]$. For this reason we refer to τ as the *trace* on $\mathcal{B}(\mathcal{H})$. It is easily verified that $\tau(AB) = \tau(BA)$ for all A and B in $\mathcal{B}(\mathcal{H})$, and that $\tau(A^*A) > 0$ if $A \neq 0$.

It is not difficult to show that, if $\mathcal{H}$ is an infinite-dimensional Hilbert space, then there is no linear functional τ on $\mathcal{B}(\mathcal{H})$ such that $\tau(AB) = \tau(BA)$ for all A and B in $\mathcal{B}(\mathcal{H})$ and $\tau(A^*A) > 0$ if $A \neq 0$.

There is, however, a linear functional τ on the ideal $\mathcal{C}_1$ of $\mathcal{B}(\mathcal{H})$ which satisfies these conditions whenever $A \in \mathcal{C}_1$ and $B \in \mathcal{B}(\mathcal{H})$. We shall now construct this functional and investigate its properties.

LEMMA 2.2.1. *Let $T \in \mathcal{C}_1$ and suppose that $\{\varphi_j : j \in J\}$ is an orthonormal basis in $\mathcal{H}$.*

(i) The sum $\sum_{j \in J} \langle T\varphi_j, \varphi_j \rangle$ exists, and does not depend on the particular choice of the orthonormal basis $\{\varphi_j : j \in J\}$.

(ii) If $T = T^*$, then $\sum_{j \in J} \langle T\varphi_j, \varphi_j \rangle = \sum_n \lambda_n$, where (λ_n) is the sequence of non-zero eigenvalues of T, counted according to their multiplicities.

Proof. Since $\mathcal{C}_1$ contains the self-adjoint and skew-adjoint parts of each of its members, (i) is an immediate consequence of (ii). Under the hypotheses of (ii) we may choose an orthonormal sequence (ψ_n) such that

$$Tx = \sum_n \lambda_n \langle x, \psi_n \rangle \psi_n \quad (x \in \mathcal{H}).$$

Then

$$\langle T\varphi_j, \varphi_j \rangle = \sum_n \lambda_n |\langle \varphi_j, \psi_n \rangle|^2,$$

$$\text{(1)} \qquad \sum_{j \in J} |\langle T\varphi_j, \varphi_j \rangle| \leqslant \sum_{j \in J} \sum_n |\lambda_n| \; |\langle \varphi_j, \psi_n \rangle|^2$$
$$= \sum_n |\lambda_n| \sum_{j \in J} |\langle \varphi_j, \psi_n \rangle|^2$$
$$= \sum_n |\lambda_n| \; \|\psi_n\|^2$$
$$= \sum_n |\lambda_n|,$$

and the last quantity is finite, by Lemma 2.1.2. It follows that the sum $\sum_{j \in J} \langle T\varphi_j, \varphi_j \rangle$ exists. A manipulation similar to the one just given, exploiting the finiteness of the double sum on the right-hand side of (1) to justify change in the order of summation, shows that $\sum_{j \in J} \langle T\varphi_j, \varphi_j \rangle = \sum_n \lambda_n$.

We shall eventually show that, in part (ii) of Lemma 2.2.1, the condition $T = T^*$ is redundant (cp. Theorem 3.3.13).

DEFINITION 2.2.2. The ideal $\mathcal{C}_1$ in $\mathcal{B}(\mathcal{H})$ is called the *trace class* of operators on $\mathcal{H}$. If $T \in \mathcal{C}_1$ and $\{\varphi_j : j \in J\}$ is an orthonormal basis in $\mathcal{H}$, then the *trace* of T, denoted by $\tau(T)$, is defined by the equation

$$\tau(T) = \sum_{j \in J} \langle T\varphi_j, \varphi_j \rangle.$$

Lemma 2.2.1 shows that $\tau(T)$ depends only on T (not on the choice of the orthonormal basis), and that $\tau(T)$ is the sum of the eigenvalues of T when $T = T^*$. Before establishing the main algebraic properties of τ, we require an auxiliary result.

LEMMA 2.2.3. *Each element A of $\mathcal{B}(\mathcal{H})$ is a finite linear combination of unitary operators.*

Proof. Since each element of $\mathcal{B}(\mathcal{H})$ can be expressed in terms of its self-adjoint and skew-adjoint parts, it is sufficient to consider the case in which $A = A^*$ and $\|A\| \leqslant 1$. In this case the spectrum $\sigma(A)$ of A is contained in the compact interval $[-1, 1]$, and we can define a continuous function f on $\sigma(A)$ by $f(t) = t + i\sqrt{(1-t^2)}$. Let $U = f(A)$, the operator corresponding to f in the functional calculus described in Theorem 1.7.7. Since

$$t = \tfrac{1}{2}[f(t)+\overline{f(t)}], \qquad f(t)\overline{f(t)} = \overline{f(t)}f(t) = 1,$$

it follows from Theorem 1.7.7. that $A = \frac{1}{2}(U+U^*)$ and $UU^* = U^*U = I$.

THEOREM 2.2.4. *Suppose that* S, $T \in \mathcal{C}_1$, $A \in \mathcal{B}(\mathcal{H})$ *and* α, β *are scalars.*

(i) $\tau(\alpha S+\beta T) = \alpha\tau(S)+\beta\tau(T)$.

(ii) $\tau(S^*) = \overline{\tau(S)}$.

(iii) $\tau(S) > 0$ if $S > 0$.

(iv) $\tau(AS) = \tau(SA)$.

Proof. Let $\{\varphi_j : j \in J\}$ be an orthonormal basis of $\mathcal{H}$. The first two parts of the theorem are obvious consequences of Definition 2.2.2. Suppose now that $S \geqslant 0$. Then $\tau(S) = \Sigma \langle S\varphi_j, \varphi_j\rangle \geqslant 0$, since each term in the summation is non-negative. If $\tau(S) = 0$ then, for each j, $0 = \langle S\varphi_j, \varphi_j\rangle = ||S^{1/2}\varphi_j||^2$, so $S^{1/2} = 0$ and $S = 0$. This proves (iii).

For (iv), it is sufficient, by Lemma 2.2.3 and the linearity of τ, to consider the case in which A is a unitary operator on $\mathcal{H}$. In this case, $\{A\varphi_j : j \in J\}$ is an orthonormal basis of $\mathcal{H}$, and

$$\begin{aligned}\tau(SA) &= \Sigma \langle SA\varphi_j, \varphi_j\rangle \\ &= \Sigma \langle SA\varphi_j, A^*A\varphi_j\rangle \\ &= \Sigma \langle AS(A\varphi_j), A\varphi_j\rangle = \tau(AS).\end{aligned}$$

2.3. *The Banach space* $\mathcal{C}_p$

In this section we introduce a norm on $\mathcal{C}_p$, and prove that $\mathcal{C}_p$ is a Banach space with properties similar to those of the sequence space ℓ_p. The main results are contained in Theorems 2.3.8, 2.3.10 and 2.3.12.

Suppose that T is a compact linear operator acting on $\mathcal{H}$, and denote by V_TH_T the polar decomposition of T. Then $T = V_TH_T$ and

$H_T = V_T^*T = (T^*T)^{1/2}$. It follows from Theorem 1.9.3, together with its proof, that there exist a decreasing sequence (μ_n) of positive real numbers (the eigenvalues of H_T, counted according to their multiplicities), and orthonormal sequences (φ_n), (ψ_n), such that

$$H_Tx = \sum_n \mu_n \langle x, \varphi_n\rangle \varphi_n, \tag{1}$$

$$Tx = \sum_n \mu_n \langle x, \varphi_n\rangle \psi_n \quad (x \in \mathcal{H}). \tag{2}$$

The sequences (μ_n), (φ_n), (ψ_n) all terminate after just k terms if T has finite rank k, and are infinite if T is not of finite rank. Given $p(\geqslant 1)$ the function f_p defined by $f_p(t) = t^p$ is continuous on the non-negative real axis, and hence also on the spectrum of the positive operator H_T. The operator $f_p(H_T)$ will be denoted by H_T^p. It follows from equation 1.9(10), and the discussion which follows it, that H_T^p is compact and that

$$H_T^px = \sum_n \mu_n^p \langle x, \varphi_n\rangle \varphi_n \quad (x \in \mathcal{H}). \tag{3}$$

It is easily verified that positive powers of H_T satisfy the usual indicial laws.

LEMMA 2.3.1. *Suppose* $1 \leqslant q \leqslant p < \infty$ *and* $T \in \mathcal{B}(\mathcal{H})$. *Then the following three conditions are equivalent.*

(i) $T \in \mathcal{C}_p$, *(ii)* $H_T \in \mathcal{C}_p$, *(iii)* $H_T^{p/q} \in \mathcal{C}_q$.

Proof. Since $\mathcal{C}_p$ is an ideal in $\mathcal{B}(\mathcal{H})$, the equivalence of (i) and (ii) follows at once from the equations $T = V_TH_T$, $H_T = V_T^*T$.

If either (ii) or (iii) is satisfied, then H_T $(= [H_T^{p/q}]^{q/p})$ is compact by Theorem 2.1.6, and we may suppose that H_T is represented as in equation (1). By Lemma 2.1.2, $H_T \in \mathcal{C}_p$ if and only if $\Sigma\, \mu_n^p < \infty$; this is equivalent to $\Sigma\, (\mu_n^{p/q})^q < \infty$, which (again

by Lemma 2.1.2) is a necessary and sufficient condition that $H_T^{p/q} \in \mathcal{C}_q$. Hence (ii) and (iii) are equivalent.

From Lemma 2.1.2 and the equivalence of conditions (i) and (ii) in the lemma just proved, it follows that a compact operator T on $\mathcal{H}$ lies in $\mathcal{C}_p$ if and only if the sequence (μ_n) of non-zero eigenvalues of $H_T = (T^*T)^{1/2}$ satisfies $\Sigma\, \mu_n^p < \infty$. This is often used as the *definition* of $\mathcal{C}_p$ (see, for example, [16]).

DEFINITION 2.3.2. Suppose that $1 \leqslant p < \infty$ and $T \in \mathcal{C}_p$. *Then* $||T||_p = [\tau(H_T^p)]^{1/p}$.

We shall adopt the convention that $||\cdot||_\infty$ is the usual norm on $\mathcal{B}(\mathcal{H})$ $(= \mathcal{C}_\infty)$. The above definition of $||T||_p$ with $1 \leqslant p < \infty$ is meaningful since, by Lemma 2.3.1, $H_T^p \in \mathcal{C}_1$ whenever $T \in \mathcal{C}_p$. It is not immediately obvious that $||\cdot||_p$ is a norm on $\mathcal{C}_p$, and before proving this we shall need some auxiliary results.

Note first that, if $1 \leqslant p < \infty$, $T \in \mathcal{C}_p$ and equation (1) holds, then H_T^p is represented in the form (3), and the remarks following Definition 2.2.2 imply that $\tau(H_T^p) = \Sigma\, \mu_n^p$. Hence

(4) $$||T||_p = [\Sigma\, \mu_n^p]^{1/p},$$

where (μ_n) is the sequence of non-zero eigenvalues of H_T, counted according to their multiplicities. Since $||T|| = ||H_T|| = \mu_1$,

(5) $$||T|| \leqslant ||T||_p \quad (T \in \mathcal{C}_p).$$

If T is a compact normal operator, and (λ_n) is the sequence of non-zero eigenvalues of T (arranged in order of decreasing magnitude and counted according to their multiplicities) then there is an orthonormal sequence (φ_n) such that

$$Tx = \sum_n \lambda_n \langle x, \varphi_n\rangle \varphi_n \quad (x \in \mathcal{H}).$$

Since $T\varphi_n = \lambda\varphi_n$, we have $T^*\varphi_n = \overline{\lambda}_n\varphi_n$ and thus

$$T^*Tx = \sum_n |\lambda_n|^2 \langle x, \varphi_n\rangle \varphi_n,$$

$$H_Tx = \sum_n |\lambda_n| \langle x, \varphi_n\rangle \varphi_n \quad (x \in \mathcal{H}).$$

Thus the sequence of non-zero eigenvalues of H_T is $(|\lambda_n|)$, and so $T \in \mathcal{C}_p$ if and only if $\Sigma\, |\lambda_n|^p < \infty$. When this is so,

$$||T||_p = [\sum_n |\lambda_n|^p]^{1/p}.$$

LEMMA 2.3.3. *For each T in $\mathcal{C}_1$, $|\tau(T)| \leqslant ||T||_1$.*

Proof. We may suppose that H_T and T are represented as in equations (1) and (2). Then $\mu_n > 0$ and $\Sigma\, \mu_n = \tau(H_T) = ||T||_1$. Let $\{\theta_j : j \in J\}$ be an orthonormal basis of $\mathcal{H}$ which contains each φ_n. If $j \in J$ and θ_j is not one of the φ_n's, then $\langle \theta_j, \varphi_n\rangle = 0$ for each n and, by (2), $T\theta_j = 0$. Thus

$$\begin{aligned} |\tau(T)| &= |\Sigma \langle T\theta_j, \theta_j\rangle| \\ &\leqslant \Sigma\, |\langle T\theta_j, \theta_j\rangle| \\ &= \Sigma\, |\langle T\varphi_n, \varphi_n\rangle| \\ &= \Sigma\, \mu_n|\langle \psi_n, \varphi_n\rangle| \\ &\leqslant \Sigma\, |\mu_n| \\ &= ||T||_1. \end{aligned}$$

LEMMA 2.3.4. *Suppose $1 \leqslant p < \infty$ and $T \in \mathcal{B}(\mathcal{H})$. Then $T \in \mathcal{C}_p$ if and only if there is a constant M such that*

(6) $$\left[\sum_{j=1}^{n} |\langle T\varphi_j, \psi_j\rangle|^p\right]^{1/p} \leqslant M$$

for every pair $(\varphi_1, \ldots, \varphi_n)$, $(\psi_1, \ldots, \psi_n)$ *of finite orthonormal systems in* $\mathcal{H}$. *When this condition is satisfied the least such constant* M *is* $||T||_p$.

Proof. Suppose that there is a constant M with the stated properties. Then $\sum_{k \in K} |\langle T\theta_k, \theta_k\rangle|^p \leqslant M^p < \infty$ for each orthonormal system $\{\theta_k : k \in K\}$ in $\mathcal{H}$, because condition (6) shows that an arbitrary finite subsum is not greater than M^p. It follows from Definition 2.1.1 that $T \in \mathcal{C}_p$.

Suppose conversely that $T \in \mathcal{C}_p$ and that T has polar decomposition $V_T H_T$. Let $(\varphi_1, \ldots, \varphi_n)$ and $(\psi_1, \ldots, \psi_n)$ be orthonormal systems in $\mathcal{H}$. Our first step is to show that, for the purpose of the present argument, the partial isometry V_T can be replaced by a unitary operator. For this, note that the vectors $H_T\varphi_1, \ldots, H_T\varphi_n$ span a finite-dimensional subspace $\mathcal{L}$ of the closed range space $\mathcal{R}(H_T)$ and that V_T maps $\mathcal{L}$ isometrically onto a subspace $\mathcal{M}$ of $\mathcal{R}(T)$. With $x_1, \ldots, x_k$ an orthonormal basis of $\mathcal{L}$ and $y_j = V_T x_j$, $y_1, \ldots, y_k$ is an orthonormal basis of $\mathcal{M}$. We may extend to form two orthonormal bases $x_1, \ldots, x_k, \ldots, x_m$ and $y_1, \ldots, y_k, \ldots, y_m$ of the subspace $\mathcal{N} = \mathcal{L} + \mathcal{M}$. Each element of $\mathcal{H}$ can be uniquely expressed in the form $a_1x_1 + \ldots + a_mx_m + z$, with z in $\mathcal{N}^\perp$, and the equation

$$U(a_1x_1 + \ldots + a_mx_m + z) = a_1y_1 + \ldots + a_my_m + z$$

defines a unitary operator U on $\mathcal{H}$. Since $Ux_j = y_j = V_Tx_j$ $(j = 1, \ldots, k)$, U and V_T coincide in their action on $\mathcal{L}$, and $UH_T\varphi_j = V_TH_T\varphi_j = T\varphi_j$ $(j = 1, \ldots, n)$. Thus

$$\begin{aligned}
\sum_{j=1}^{n} |\langle T\varphi_j, \psi_j\rangle|^p &= \Sigma\, |\langle UH_T\varphi_j, \psi_j\rangle|^p \\
&= \Sigma\, |\langle H_T^{1/2}\varphi_j, H_T^{1/2}U^*\psi_j\rangle|^p \\
&\leqslant \Sigma\, ||H_T^{1/2}\varphi_j||^p\, ||H_T^{1/2}U^*\psi_j||^p \\
&\leqslant [\Sigma\, ||H_T^{1/2}\varphi_j||^{2p}]^{1/2}[\Sigma\, ||H_T^{1/2}U^*\psi_j||^{2p}]^{1/2} \\
&= [\Sigma\, \langle H_T\varphi_j, \varphi_j\rangle^p]^{1/2} \times \\
&\qquad [\Sigma\, \langle H_TU^*\psi_j, U^*\psi_j\rangle^p]^{1/2}.
\end{aligned}$$

Since $U^*\psi_1, \ldots, U^*\psi_n$ is an orthonormal system in $\mathcal{H}$, it follows, from Lemma 2.1.2(ii) and the expression (4) for $||T||_p$ in terms of the eigenvalues of H_T, that the final quantity in the above series of inequalities does not exceed $||T||_p^p$. Hence condition (6) is satisfied, with $M = ||T||_p$.

Finally we show that, if M is any constant which satisfies (6), then $M \geqslant ||T||_p$. We may suppose that T is represented as in equation (2). Then (φ_n) and (ψ_n) are orthonormal sequences and $T\varphi_n = \mu_n\psi_n$. For each m we have

$$M \geqslant [\sum_{n=1}^{m} |\langle T\varphi_n, \psi_n\rangle|^p]^{1/p} = [\sum_{n=1}^{m} \mu_n^p]^{1/p}.$$

Hence $$M \geqslant [\sum_n \mu_n^p]^{1/p} = ||T||_p.$$

COROLLARY 2.3.5. *For each* T *in* $\mathcal{C}_p$, $||T^*||_p = ||T||_p$.

COROLLARY 2.3.6. *Suppose that* $1 \leqslant p < \infty$, $T \in \mathcal{C}_p$, *and* (λ_n) *is the sequence of non-zero eigenvalues of* T, *counted according to their algebraic multiplicities. Then*

$$[\sum_n |\lambda_n|^p]^{1/p} \leqslant ||T||_p.$$

Proof. By Theorem 1.8.6 there is an orthonormal sequence (φ_n) in $\mathcal{H}$ for which $\langle T\varphi_n, \varphi_n\rangle = \lambda_n$; and

$$[\sum_n |\langle T\varphi_n, \varphi_n\rangle|^p]^{1/p} \leqslant ||T||_p$$

since, by Lemma 2.3.4, each finite subsum is not greater than $||T||_p$.

Let $\mathcal{F}$ be the set of all operators of finite rank on $\mathcal{H}$. When $x, y \in \mathcal{H}$, we define $x \otimes y$ in $\mathcal{F}$ (as at the end of §1.9) by $(x \otimes y)z = \langle z, y\rangle x$ $(z \in \mathcal{H})$. If $\{\varphi_j : j \in J\}$ is an orthonormal basis of $\mathcal{H}$,

$$\begin{aligned}\tau(x \otimes y) &= \sum_j \langle (x \otimes y)\varphi_j, \varphi_j\rangle\\ &= \sum_j \langle \varphi_j, y\rangle\langle x, \varphi_j\rangle\\ &= \langle x, y\rangle,\end{aligned}$$

by Parseval's equation. We claim also that

$$||x \otimes y||_p = ||x||\ ||y|| \quad (1 \leqslant p \leqslant \infty).$$

For this, it is sufficient to consider the case in which $||x|| = ||y|| = 1$. In that case, the discussion at the end of §1.9 shows that, if $x \otimes y$ has polar decomposition VH, then $H = y \otimes y$. Hence the only eigenvalue of H is 1, and its multiplicity is 1. By (4), $||x \otimes y||_p = 1$ $(1 \leqslant p \leqslant \infty)$.

We recall from §1.3 that, if $1 \leqslant p \leqslant \infty$, then the index q conjugate to p is defined by $p^{-1}+q^{-1} = 1$, with the obvious interpretation when p is 1 or ∞.

LEMMA 2.3.7. *Suppose that $1 \leqslant p \leqslant \infty$, q is the index conjugate to p, and $T \in \mathcal{B}(\mathcal{H})$. Then $T \in \mathcal{C}_p$ if and only if*

$$\sup\{|\tau(FT)| : F \in \mathcal{F},\ ||F||_q \leqslant 1\} < \infty. \tag{7}$$

When this is so, the value of the supremum is $||T||_p$.

Proof. Suppose first that the supremum in (7) is a finite quantity M. We shall prove that $T \in \mathcal{C}_p$ and $||T||_p \leqslant M$. In the case $p = \infty$, it is given that $T \in \mathcal{B}(\mathcal{H}) = \mathcal{C}_\infty$; and, if x and y are unit vectors in $\mathcal{H}$, then $x \otimes y \in \mathcal{F}$, $||x \otimes y||_1 = 1$, hence

$$\begin{aligned}M \geqslant |\tau((x \otimes y)T)| &= |\tau(x \otimes T^*y)|\\ &= |\langle x, T^*y\rangle|\\ &= |\langle Tx, y\rangle|.\end{aligned}$$

By taking the supremum of the right-hand side over all unit vectors x and y, we obtain $M \geqslant ||T|| = ||T||_\infty$. We now turn to the case in which $1 \leqslant p < \infty$. Let $(\varphi_1, \ldots, \varphi_n)$ and $(\psi_1, \ldots, \psi_n)$ be orthonormal systems in $\mathcal{H}$ and let $\alpha_k = \langle T\varphi_k, \psi_k\rangle$. We shall prove that $\Sigma |\alpha_k|^p \leqslant M^p$, whereupon Lemma 2.3.4 shows that $T \in \mathcal{C}_p$ and $||T||_p \leqslant M$, as required. By deleting any φ_k and ψ_k for which $\langle T\varphi_k, \psi_k\rangle = 0$, we may suppose that each α_k is non-zero and define $\beta_k = \alpha_k^{-1}|\alpha_k|$. Let

$$Bx = \sum_{k=1}^{n} |\alpha_k|^{p-1}\langle x, \psi_k\rangle\beta_k\varphi_k,$$

$$Hx = \sum_{k=1}^{n} |\alpha_k|^{p-1}\langle x, \psi_k\rangle\psi_k \quad (x \in \mathcal{H}).$$

Then $B \in \mathcal{F}$, and it follows from the discussion at the end of §1.9 that B has polar decomposition of the form VH. Since $(p-1)q = p$ if $1 < p < \infty$ and the eigenvalues of H are $|\alpha_1|^{p-1}, \ldots, |\alpha_n|^{p-1}$, it follows from equation (4) that

$$(8)\qquad \|B\|_q = \begin{cases} [\sum_{k=1}^{n} |\alpha_k|^p]^{1/q} & \text{if } 1 < p < \infty, \\ 1 & \text{if } p = 1. \end{cases}$$

For each x in $\mathcal{H}$,

$$\begin{aligned} BTx &= \Sigma\, |\alpha_k|^{p-1} \langle Tx, \psi_k \rangle \beta_k \varphi_k \\ &= \Sigma\, |\alpha_k|^{p-1} \beta_k \langle x, T^*\psi_k \rangle \varphi_k, \end{aligned}$$

so $BT = \Sigma\, |\alpha_k|^{p-1}\beta_k(\varphi_k \otimes T^*\psi_k)$ and

$$\begin{aligned} \tau(BT) &= \Sigma\, |\alpha_k|^{p-1}\beta_k \langle \varphi_k, T^*\psi_k \rangle \\ &= \Sigma\, |\alpha_k|^{p-1}\beta_k \alpha_k \\ &= \Sigma\, |\alpha_k|^p. \end{aligned}$$

With $F = \|B\|_q^{-1}B$, we have $F \in \mathcal{F}$ and $\|F\|_q = 1$. Since M is the supremum in (7), it follows from (8) that

$$M \geqslant |\tau(FT)| = [\sum_{k=1}^{n} |\alpha_k|^p]^{1/p}.$$

This completes the proof that, if the supremum in (7) is a finite quantity M, then $T \in \mathcal{C}_p$ and $\|T\|_p \leqslant M$.

Suppose conversely that $T \in \mathcal{C}_p$. We shall prove that the supremum in (7) is finite and not greater than $\|T\|_p$. For this purpose, suppose that $F \in \mathcal{F}$ and $\|F\|_q \leqslant 1$. If VH is the polar decomposition of F, there exist positive $\alpha_1, \ldots, \alpha_k$ and orthonormal systems $(\varphi_1, \ldots, \varphi_k)$ and $(\psi_1, \ldots, \psi_k)$ such that

$$Hx = \sum_{k=1}^{n} \alpha_k \langle x, \varphi_k \rangle \varphi_k,$$

$$Fx = \sum_{k=1}^{n} \alpha_k \langle x, \varphi_k \rangle \psi_k \quad (x \in \mathcal{H}).$$

Since $\|F\|_q \leqslant 1$, it follows that

$$(9)\qquad \begin{cases} \sum_{k=1}^{n} \alpha_k^q \leqslant 1 & \text{if } 1 \leqslant q < \infty, \\ \alpha_k \leqslant 1 \quad (k = 1, \ldots, n) & \text{if } q = \infty. \end{cases}$$

For each x in $\mathcal{H}$,

$$\begin{aligned} FTx &= \sum_{k=1}^{n} \alpha_k \langle Tx, \varphi_k \rangle \psi_k \\ &= \sum_{k=1}^{n} \alpha_k \langle x, T^*\varphi_k \rangle \psi_k, \end{aligned}$$

so $FT = \Sigma\, \alpha_k(\psi_k \otimes T^*\varphi_k)$ and

$$\tau(FT) = \Sigma\, \alpha_k \langle \psi_k, T^*\varphi_k \rangle = \Sigma\, \alpha_k \beta_k,$$

where $\beta_k = \langle T\psi_k, \varphi_k \rangle$. By Lemma 2.3.4,

$$(10)\qquad \begin{cases} [\sum_{k=1}^{n} |\beta_k|^p]^{1/p} \leqslant \|T\|_p & \text{if } 1 \leqslant p < \infty, \\ |\beta_k| \leqslant \|T\|_\infty \quad (k = 1, \ldots, n) & \text{if } p = \infty. \end{cases}$$

From (9) and (10) it follows, by use of Hölder's inequality, that

$$|\tau(FT)| = |\sum_{k=1}^{n} \alpha_k \beta_k| \leqslant \|T\|_p.$$

We deduce that, if $T \in \mathcal{C}_p$, then the supremum in (7) is finite and not greater than $\|T\|_p$. This, together with our previous result in the opposite direction, completes the proof of the lemma.

THEOREM 2.3.8. *Suppose that* $1 \leqslant p < \infty$. *Then* $\|\cdot\|_p$ *is a norm on* $\mathcal{C}_p$, *and with this norm* $\mathcal{C}_p$ *is a Banach space. The set* $\mathcal{F}$ *of all operators of finite rank on* $\mathcal{H}$ *is an everywhere dense linear subspace of* $\mathcal{C}_p$.

Proof. Since τ is a linear functional on $\mathcal{C}_1$ and, by Lemma 2.3.7,

$$||T||_p = \sup\{|\tau(FT)| : F \in \mathcal{F},\ ||F||_q \leqslant 1\}$$

for each T in $\mathcal{C}_p$, it follows that

$$||T||_p \geqslant 0, \qquad ||\alpha T||_p = |\alpha|\ ||T||_p, \qquad ||S+T||_p \leqslant ||S||_p + ||T||_p$$

for each S and T in $\mathcal{C}_p$ and each complex number α. If $||T||_p = 0$ then by (5) $||T|| = 0$, and so $T = 0$. This shows that $||\cdot||_p$ is a norm on $\mathcal{C}_p$.

Suppose that (T_p) is a Cauchy sequence in $\mathcal{C}_p$. Given a positive ϵ there is an integer $N(\epsilon)$ such that

$$(11) \qquad ||T_m - T_n||_p < \epsilon \quad (m,\ n \geqslant N(\epsilon)).$$

Hence $||T_m - T_n|| < \epsilon$ whenever $m,\ n \geqslant N(\epsilon)$, so (T_n) is a Cauchy sequence in $\mathcal{B}(\mathcal{H})$ and converges, with respect to the norm on $\mathcal{B}(\mathcal{H})$, to some T in $\mathcal{B}(\mathcal{H})$. If $x,\ y \in \mathcal{H}$ then

$$\begin{aligned} \tau((x \otimes y)T_n) &= \tau(x \otimes T_n^* y) \\ &= \langle T_n x,\ y \rangle \\ &\to \langle Tx,\ y \rangle = \tau((x \otimes y)T) \end{aligned}$$

as $n \to \infty$. It follows, from the linearity of τ, that $\tau(FT_n) \to \tau(FT)$ as $n \to \infty$, for each F in $\mathcal{F}$. With F in $\mathcal{F}$ and $||F||_q \leqslant 1$, Lemma 2.3.7 and (11) imply that

$$|\tau(FT_m - FT_n)| < \epsilon \quad (m,\ n \geqslant N(\epsilon)),$$

and by taking the limit of the left hand side as $n \to \infty$ we obtain

$$|\tau(FT_m - FT)| \leqslant \epsilon \quad (m \geqslant N(\epsilon)).$$

Hence

$$\sup\{|\tau(FT_m - FT)| : F \in \mathcal{F},\ ||F||_q \leqslant 1\} \leqslant \epsilon$$

when $m \geqslant N(\epsilon)$, and Lemma 2.3.7 shows that

$$T_m - T \in \mathcal{C}_p, \qquad ||T_m - T||_p \leqslant \epsilon \quad (m \geqslant N(\epsilon)).$$

It follows that $T \in \mathcal{C}_p$ and that (T_n) converges to T with respect to the norm on $\mathcal{C}_p$. This proves that $\mathcal{C}_p$ is complete.

We show finally that each T in $\mathcal{C}_p$ lies in the closure of $\mathcal{F}$. We may suppose that T is represented as in (2), the summation being over all positive integers. For $m = 1, 2, \ldots$, define T_m in $\mathcal{F}$ by

$$(12) \qquad T_m x = \sum_{n=1}^{m} \mu_n \langle x,\ \varphi_n \rangle \psi_n \quad (x \in \mathcal{H}).$$

Then

$$(T - T_m)x = \sum_{n=m+1}^{\infty} \mu_n \langle x,\ \varphi_n \rangle \psi_n \quad (x \in \mathcal{H}).$$

By applying (4), with $T - T_m$ in place of T, we obtain

$$||T - T_m||_p = \Big[\sum_{n=m+1}^{\infty} \mu_n^p\Big]^{1/p} \to 0$$

as $m \to \infty$. Hence T lies in the closure of $\mathcal{F}$.

Our next objective is to prove a generalization of Hölder's inequality (Theorem 2.3.10). The following lemma is subsumed in Theorem 2.3.10 but is needed in its proof.

LEMMA 2.3.9. *Suppose that* $2 \leqslant r,\ s \leqslant \infty$, $t^{-1} = r^{-1} + s^{-1}$, $R \in \mathcal{C}_r$ *and* $S \in \mathcal{C}_s$. *Then* $RS \in \mathcal{C}_t$ *and* $||RS||_t \leqslant ||R||_r ||S||_s$.

Proof. If $r = s = \infty$ then $t = \infty$ and the result is trivial, so we shall assume that at least one of r and s is finite and hence that t is finite. Let $a = rt^{-1}$, $b = st^{-1}$, and note that a and b are conjugate indices. We give the proof for the case in which $1 < a < \infty$, and leave the reader to supply the simple modifications required when $a = 1$ or ∞. Given any finite orthonormal systems $(\varphi_1, \ldots, \varphi_n)$ and $(\psi_1, \ldots, \psi_n)$ in $\mathcal{H}$, we have

$$\begin{aligned}\sum_{j=1}^{n} |\langle RS\varphi_j, \psi_j\rangle|^t &= \sum_{j=1}^{n} |\langle S\varphi_j, R^*\psi_j\rangle|^t \\ &\leqslant \sum_{j=1}^{n} ||S\varphi_j||^t ||R^*\psi_j||^t \\ &\leqslant [\sum_{j=1}^{n} ||S\varphi_j||^{bt}]^{1/b} [\sum_{j=1}^{n} ||R^*\psi_j||^{at}]^{1/a} \\ &= [\sum_{j=1}^{n} \langle S^*S\varphi_j, \varphi_j\rangle^{s/2}]^{1/b} \\ &\quad \times [\sum_{j=1}^{n} \langle RR^*\psi_j, \psi_j\rangle^{r/2}]^{1/a}\end{aligned}$$

Hence

$$\begin{aligned}[\sum_{j=1}^{n} |\langle RS\varphi_j, \psi_j\rangle|^t]^{1/t} &\leqslant [\sum_{j=1}^{n} \langle S^*S\varphi_j, \varphi_j\rangle^{s/2}]^{1/s} \\ &\quad \times [\sum_{j=1}^{n} \langle RR^*\psi_j, \psi_j\rangle^{r/2}]^{1/r}.\end{aligned} \tag{13}$$

Now suppose that (μ_n) is the sequence of non-zero eigenvalues of $(S^*S)^{1/2}$, counted according to their multiplicities, and note that (μ_n^2) is the corresponding sequence for S^*S. Since $s/2 \geqslant 1$ and, by (4), $\Sigma(\mu_n^2)^{s/2} = ||S||_s^s$, it follows from Lemma 2.1.2 that

$$\sum_{j=1}^{n} \langle S^*S\varphi_j, \varphi_j\rangle^{s/2} \leqslant ||S||_s^s;$$

and a similar argument shows that

$$\sum_{j=1}^{n} \langle RR^*\psi_j, \psi_j\rangle^{r/2} \leqslant ||R^*||_r^r = ||R||_r^r.$$

These inequalities, together with (13), yield

$$[\sum_{j=1}^{n} |\langle RS\varphi_j, \psi_j\rangle|^t]^{1/t} \leqslant ||R||_r ||S||_s.$$

Since this has been proved for any finite orthonormal systems $(\varphi_1, \ldots, \varphi_n)$ and $(\psi_1, \ldots, \psi_n)$ in $\mathcal{H}$, it follows from Lemma 2.3.4 that $RS \in \mathcal{C}_t$ and $||RS||_t \leqslant ||R||_r ||S||_s$.

THEOREM 2.3.10. *Suppose that $1 \leqslant r, s, t \leqslant \infty$, $t^{-1} = r^{-1}+s^{-1}$, $R \in \mathcal{C}_r$ and $S \in \mathcal{C}_s$. Then $RS \in \mathcal{C}_t$ and $||RS||_t \leqslant ||R||_r ||S||_s$.*

Proof. We deal separately with three cases which together cover all possibilities.

Case 1. $r, s \geqslant 2$. In this case the result has already been proved in Lemma 2.3.9.

Case 2. $1 \leqslant s < 2$. Suppose first that $S \in \mathcal{F}$. Let p and q be the indices conjugate to s and t, respectively; note that $q^{-1}+r^{-1}+s^{-1} = 1$, whence $q, r > 2$ and $q^{-1}+r^{-1} = p^{-1}$. Suppose that $F \in \mathcal{F}$ and $||F||_q \leqslant 1$. By using Theorem 2.2.4 (iv) (with FR in place of A), Lemma 2.3.7 (with s in place of q) and Lemma 2.3.9 (with q, r and p in place of r, s and t, respectively) we obtain

$$\begin{aligned}|\tau(FRS)| = |\tau(SFR)| &\leqslant ||S||_s ||FR||_p \\ &\leqslant ||S||_s ||F||_q ||R||_r \leqslant ||R||_r ||S||_s.\end{aligned}$$

Since this inequality is satisfied whenever $F \in \mathcal{F}$ and $||F||_q \leqslant 1$, it follows from Lemma 2.3.7 (with t in place of p) that $RS \in \mathcal{C}_t$ and $||RS||_t \leqslant ||R||_r ||S||_s$.

We have proved the theorem in case 2, subject to the additional assumption that $S \in \mathcal{F}$. We continue considering case 2, but now with a general S in $\mathcal{C}_s$. By Theorem 2.3.8 there is a sequence (S_n) in $\mathcal{F}$ for which $||S-S_n||_s \to 0$. The result just proved shows that $RS_n \in \mathcal{C}_t$, $||RS_n||_t \leqslant ||R||_r ||S_n||_s$ and

$$||RS_m - RS_n||_t \leqslant ||R||_r ||S_m - S_n||_s \to 0$$

as min $(m, n) \to \infty$. Since $\mathcal{C}_t$ is complete, the sequence (RS_n) converges with respect to $||\cdot||_t$ to some T in $\mathcal{C}_t$. We have

$$||RS_n - T|| \leqslant ||RS_n - T||_t \to 0,$$

$$||RS_n - RS|| \leqslant ||R|| \ ||S_n - S|| \leqslant ||R|| \ ||S_n - S||_s \to 0;$$

so $RS = T \in \mathcal{C}_t$, and

$$\begin{aligned} ||RS||_t &= ||T||_t = \lim ||RS_n||_t \\ &\leqslant ||R||_r \lim ||S_n||_s = ||R||_r ||S||_s. \end{aligned}$$

This completes the proof of the theorem in case 2.

Case 3. $1 \leqslant r < 2$. Since $S^* \in \mathcal{C}_s$ and $R^* \in \mathcal{C}_r$ we may apply the known result in case 2 (with r and s reversed) to obtain $S^*R^* \in \mathcal{C}_t$ and $||S^*R^*||_t \leqslant ||S^*||_s ||R^*||_r$. Hence $RS \in \mathcal{C}_t$ and, by Corollary 2.3.5, $||RS||_t \leqslant ||R||_r ||S||_s$.

This completes the proof of Theorem 2.3.10.

COROLLARY 2.3.11. *Suppose that $A, B \in \mathcal{B}(\mathcal{H})$ and $S \in \mathcal{C}_1$. Then $||ASB||_1 \leqslant ||A|| \ ||S||_1 ||B||$.*

Proof. By Theorem 2.3.10,

$$\begin{aligned} ||ASB||_1 &\leqslant ||A||_\infty ||SB||_1 \\ &\leqslant ||A||_\infty ||S||_1 ||B||_\infty = ||A|| \ ||S||_1 ||B||. \end{aligned}$$

We now identify the dual space of the Banach space $\mathcal{C}_p$, where $1 \leqslant p < \infty$.

THEOREM 2.3.12. *Suppose that $1 \leqslant p < \infty$ and q is the index conjugate to p. Then for each T in $\mathcal{C}_q$ the equation $f_T(S) = \tau(ST)$ $(S \in \mathcal{C}_p)$ defines a continuous linear functional on $\mathcal{C}_p$. Furthermore the mapping $T \to f_T$ is an isometric isomorphism from $\mathcal{C}_q$ onto the dual space $(\mathcal{C}_p)^*$.*

Proof. Suppose that $S \in \mathcal{C}_p$ and $T \in \mathcal{C}_q$. By Theorem 2.3.10 and Lemma 2.3.3 we have $ST \in \mathcal{C}_1$ and

$$|\tau(ST)| \leqslant ||ST||_1 \leqslant ||S||_p ||T||_q.$$

It follows that, for each T in $\mathcal{C}_q$, the equation $f_T(S) = \tau(ST)$ $(S \in \mathcal{C}_p)$ defines a continuous linear functional f_T on $\mathcal{C}_p$, with $||f_T|| \leqslant ||T||_q$. By Lemma 2.3.7 we have

$$\begin{aligned} ||f_T|| &\geqslant \sup \{|f_T(F)| : F \in \mathcal{F}, ||F||_p \leqslant 1\} \\ &= \sup \{|\tau(FT)| : F \in \mathcal{F}, ||F||_p \leqslant 1\} \\ &= ||T||_q, \end{aligned}$$

so $||f_T|| = ||T||_q$. Since the mapping $T \to f_T$ from $\mathcal{C}_q$ into $(\mathcal{C}_p)^*$ is obviously linear, it remains only to prove that its range is the whole of $(\mathcal{C}_p)^*$.

Let f be a continuous linear functional on $\mathcal{C}_p$, and (using the notation introduced at the end of §1.9) define $L(x, y) = f(x \otimes y)$ for each x and y in $\mathcal{H}$. Since $L(x, y)$ is linear in x, conjugate linear in y

and satisfies

$$|L(x, y)| = |f(x \otimes y)| \leqslant ||f||\,||x \otimes y||_p = ||f||\,||x||\,||y||,$$

it follows from Theorem 1.7.1 that there is an element T of $\mathcal{B}(\mathcal{H})$ such that $L(x, y) = \langle Tx, y\rangle$. Hence

$$\begin{aligned} f(x \otimes y) = \langle Tx, y\rangle &= \langle x, T^*y\rangle \\ &= \tau(x \otimes T^*y) = \tau((x \otimes y)T), \end{aligned}$$

and so $f(F) = \tau(FT)$ for each F in $\mathcal{F}$. Thus

$$\sup \{|\tau(FT)| : F \in \mathcal{F},\ ||F||_p \leqslant 1\} \leqslant ||f|| < \infty,$$

and by Lemma 2.3.7 we have $T \in \mathcal{C}_q$. Finally, $f(F) = \tau(FT) = f_T(F)$ for each F in $\mathcal{F}$ and, since $\mathcal{F}$ is dense in $\mathcal{C}_p$, it follows by continuity of f and f_T that $f = f_T$. This completes the proof of the theorem.

It follows easily from Theorem 2.3.12 that the Banach space $\mathcal{C}_p$ is reflexive if $1 < p < \infty$. For $(\mathcal{C}_p)^*$ can be identified with $\mathcal{C}_q$ and $(\mathcal{C}_q)^*$ can be identified with $\mathcal{C}_p$; these identifications give rise to an isometric isomorphism from $\mathcal{C}_p$ onto $(\mathcal{C}_p)^{**}$, and a routine manipulation shows that this isomorphism is the canonical one described in §1.1.

THEOREM 2.3.13. *Suppose that* $1 \leqslant p \leqslant \infty$, *$q$ is the index conjugate to p,* $S \in \mathcal{C}_p$ *and* $T \in \mathcal{C}_q$. *Then* $ST,\ TS \in \mathcal{C}_1$ *and* $\tau(ST) = \tau(TS)$.

Proof. Since the roles of p and q can be reversed we may assume that $p \neq \infty$. Then $\mathcal{F}$ is dense in $\mathcal{C}_p$ and so, given any positive ϵ, there exists F in $\mathcal{F}$ for which $||S-F||_p < \epsilon$. By Theorem 2.3.10 and Lemma 2.3.3 we have $ST,\ TS \in \mathcal{C}_1$,

$$|\tau(ST-FT)| \leqslant ||ST-FT||_1 \leqslant ||S-F||_p ||T||_q \leqslant \epsilon||T||_q,$$

and similarly $|\tau(TS-TF)| \leqslant \epsilon||T||_q$. By Theorem 2.2.4 (iv) $\tau(FT) = \tau(TF)$; hence $|\tau(ST)-\tau(TS)| \leqslant 2\epsilon||T||_q$ and, since ϵ is arbitrary, $\tau(ST) = \tau(TS)$.

The most important cases of Theorem 2.3.13 are those in which $p = 1$ and $q = \infty$ (already included in Theorem 2.2.4) or $p = q = 2$.

2.4. *The Schmidt class*

DEFINITION 2.4.1. The ideal $\mathcal{C}_2$ in $\mathcal{B}(\mathcal{H})$ is called the *Schmidt class* of operators on $\mathcal{H}$.

We recall that $\mathcal{C}_2$ is a Banach space with respect to the norm $||\cdot||_2$ described in Definition 2.3.2. The following theorem shows that $\mathcal{C}_2$ is, in fact, a Hilbert space, and has an orthonormal basis consisting of operators of the form $x \otimes y$ (the notation being that introduced at the end of §1.9).

THEOREM 2.4.2. *The Schmidt class* $\mathcal{C}_2$ *is a Hilbert space with respect to the inner product* [,] *defined by* $[S, T] = \tau(T^*S)$ *for each* S *and* T *in* $\mathcal{C}_2$. *The norm derived from this inner product is* $||\cdot||_2$.

If $\{\varphi_j : j \in J\}$ *and* $\{\psi_j : j \in J\}$ *are orthonormal bases in* $\mathcal{H}$ *and* $F_{j,k} = \varphi_j \otimes \psi_k$, *then* $\{F_{j,k} : j, k \in J\}$ *is an orthonormal basis in* $\mathcal{C}_2$.

Proof. It is clear that $[S, T]$ is linear in S and that its conjugate complex number is $[T, S]$. If $V_T H_T$ is the polar decomposition of an element T of $\mathcal{C}_2$, then

$$[T, T]^{1/2} = \{\tau(T^*T)\}^{1/2} = \{\tau(H_T^2)\}^{1/2} = ||T||_2.$$

Since $||\cdot||_2$ is a norm under which $\mathcal{C}_2$ is complete, it follows that $\mathcal{C}_2$ is a Hilbert space with $[\ ,\ ]$ as inner product.

Since

$$F^*_{j,k} F_{\ell,m} = (\psi_k \otimes \varphi_j)(\varphi_\ell \otimes \psi_m) = \langle \varphi_\ell, \varphi_j \rangle (\psi_k \otimes \psi_m),$$

we have

$$[F_{\ell,m}, F_{j,k}] = \tau(F^*_{j,k} F_{\ell,m}) = \langle \varphi_\ell, \varphi_j \rangle \langle \psi_k, \psi_m \rangle = \begin{cases} 1 & \text{if } \ell = j \text{ and } m = k, \\ 0 & \text{otherwise.} \end{cases}$$

Hence $\{F_{j,k} : j, k \in J\}$ is an orthonormal system in $\mathcal{C}_2$. Note that, if $T \in \mathcal{C}_2$, then

$$[T, F_{j,k}] = \tau(F^*_{j,k} T) = \tau(\psi_k \otimes T^*\varphi_j) = \langle T\psi_k, \varphi_j \rangle.$$

If T is orthogonal to each $F_{j,k}$, then $\langle T\psi_k, \varphi_j \rangle = 0$ for all j, k in J; since T is continuous, and finite linear combinations of ψ_k's, or of φ_j's, are everywhere dense in $\mathcal{H}$, it follows that $\langle Tx, y \rangle = 0$ for each x and y in $\mathcal{H}$ and hence that $T = 0$. Thus $\{F_{j,k} : j, k \in J\}$ is an orthonormal basis in $\mathcal{C}_2$.

THEOREM 2.4.3. *Suppose that $T \in \mathcal{B}(\mathcal{H})$, $\{\varphi_j : j \in J\}$ and $\{\psi_j : j \in J\}$ are orthonormal bases in $\mathcal{H}$, and $F_{j,k} = \varphi_j \otimes \psi_k$. Then the following three conditions are equivalent.*

(i) $\sum_{j \in J} ||T\psi_j||^2 < \infty$.

(ii) $\sum_{j,k \in J} |\langle T\psi_k, \varphi_j \rangle|^2 < \infty$.

(iii) $T \in \mathcal{C}_2$.

When these conditions are satisfied, the sums occurring in (i) and (ii) are both equal to $||T||_2^2$, and

$$T = \sum_{j,k \in J} \langle T\psi_k, \varphi_j \rangle F_{j,k},$$

the summation converging with respect to $||\cdot||_2$.

Proof. Since $||T\psi_k||^2 = \sum_{j \in J} |\langle T\psi_k, \varphi_j \rangle|^2$ it follows that

(1) $$\sum_{k \in J} ||T\psi_k||^2 = \sum_{k \in J} \sum_{j \in J} |\langle T\psi_k, \varphi_j \rangle|^2$$

(these sums being finite or infinite), so (i) and (ii) are equivalent.

If $T \in \mathcal{C}_2$ then, since $\{F_{j,k} : j, k \in J\}$ is an orthonormal basis in $\mathcal{C}_2$,

(2) $$||T||_2^2 = \sum_{j,k \in J} |[T, F_{j,k}]|^2 = \sum_{j,k \in J} |\langle T\psi_k, \varphi_j \rangle|^2,$$

(3) $$T = \sum_{j,k \in J} [T, F_{j,k}] F_{j,k} = \sum_{j,k \in J} \langle T\psi_k, \varphi_j \rangle F_{j,k},$$

where the summations in (3) converge with respect to $||\cdot||_2$. In particular, it follows from (2) that (ii) is satisfied, so (iii) implies (ii).

Suppose conversely that $T \in \mathcal{B}(\mathcal{H})$ and (ii) is satisfied. With $c_{j,k} = \langle T\psi_k, \varphi_j \rangle$, we have $\sum_{j,k} |c_{j,k}|^2 < \infty$, hence $\sum_{j,k} c_{j,k} F_{j,k}$ converges with respect to $||\cdot||_2$ to an element S of $\mathcal{C}_2$ for which $[S, F_{j,k}] = c_{j,k}$. Since

$$\langle S\psi_k, \varphi_j \rangle = [S, F_{j,k}] = c_{j,k} = \langle T\psi_k, \varphi_j \rangle$$

for each j and k in J, it follows that $T = S \in \mathcal{C}_2$. Hence (ii) implies (iii).

We have now shown that conditions (i), (ii) and (iii) are equivalent and that, given these condition, equations (1), (2) and (3) are satisfied. This completes the proof of the theorem.

It follows at once from Theorem 2.4.3 that, if $T \in \mathcal{B}(\mathcal{H})$ and *either* (i) or (ii) is satisfied for *one particular* choice of the orthonormal bases involved, then *both* (i) and (ii) are satisfied for *any* choice of these orthonormal bases.

We now give a simple description of the Schmidt class of operators on an L_2 space.

THEOREM 2.4.4. *Let $\mathcal{H} = L_2(E, \mu)$ and, for each h in $L_2(E \times E, \mu \times \mu)$, let T_h be the associated integral operator on $\mathcal{H}$ defined by*

$$(T_h f)(s) = \int_E h(s, t) f(t) d\mu(t).$$

Then the mapping $h \to T_h$ is an isometric isomorphism from $L_2(E \times E, \mu \times \mu)$ onto $\mathcal{C}_2$.

Proof. Let $\{\varphi_j : j \in J\}$ be an orthonormal basis in $L_2(E, \mu)$. We may define an orthonormal system $\{\psi_{j,k} : j, k \in J\}$ in $L_2(E \times E, \mu \times \mu)$ by $\psi_{j,k}(s, t) = \varphi_j(s)\overline{\varphi_k(t)}$, and for each h in $L_2(E \times E, \mu \times \mu)$ we have

$$\begin{aligned} \langle T_h \varphi_k, \varphi_j \rangle &= \iint h(s, t) \varphi_k(t) \overline{\varphi_j(s)} d\mu(t) d\mu(s) \\ &= \langle h, \psi_{j,k} \rangle. \end{aligned}$$

It follows that, if $\langle h, \psi_{j,k} \rangle = 0$ for each j and k in J, then $T_h = 0$ and hence (see §1.8) $h(s, t) = 0$ almost everywhere on $E \times E$. Thus $\{\psi_{j,k} : j, k \in J\}$ is an orthonormal basis of $L_2(E \times E, \mu \times \mu)$ and, for each h in $L_2(E \times E, \mu \times \mu)$,

$$\begin{aligned} ||h||^2 &= \sum_{j,\, k \in J} |\langle h, \psi_{j,k} \rangle|^2 \\ &= \sum_{j,\, k \in J} |\langle T_h \varphi_k, \varphi_j \rangle|^2. \end{aligned}$$

By Theorem 2.4.3, $T_h \in \mathcal{C}_2$ and $||T_h||_2 = ||h||$. It follows that the mapping $h \to T_h$ is an isometric isomorphism from $L_2(E \times E, \mu \times \mu)$ onto a subspace $\mathfrak{M}$ of $\mathcal{C}_2$. Since $L_2(E \times E, \mu \times \mu)$ is complete, the same is true of $\mathfrak{M}$, and thus $\mathfrak{M}$ is a closed subspace of $\mathcal{C}_2$. It is easily verified that $T_h = \varphi_j \otimes \varphi_k$ when $h = \psi_{j,k}$, so $\varphi_j \otimes \varphi_k \in \mathfrak{M}$. By Theorem 2.4.2, $\{\varphi_j \otimes \varphi_k : j, k \in J\}$ is an orthonormal basis of $\mathcal{C}_2$, so $\mathfrak{M} = \mathcal{C}_2$.

We now revert to the study of the Schmidt class of operators on a general Hilbert space.

THEOREM 2.4.5. *Suppose that $T \in \mathcal{C}_2$ and (λ_n) is the sequence of non-zero eigenvalues of T, counted according to their algebraic multiplicities. Then*

$$\sum_n |\lambda_n|^2 \leqslant ||T||_2^2; \tag{4}$$

equality occurs if and only if T is normal.

Proof. The inequality (4) is a special case of Corollary 2.3.6. However, in order to determine the conditions under which equality occurs, we need a different proof of (4). By Theorem 1.8.6 there is an orthonormal sequence (ψ_n) in $\mathcal{H}$ such that $\langle T\psi_n, \psi_n \rangle = \lambda_n$. Let $\{\varphi_j : j \in J\}$ be an orthonormal basis of $\mathcal{H}$ which contains each ψ_n, and define $K = \{ j \in J : \varphi_j$ is one of the ψ_n's$\}$. Then

$$\begin{aligned} \sum_n |\lambda_n|^2 &= \sum_n |\langle T\psi_n, \psi_n \rangle|^2 \\ &= \sum_{k \in K} |\langle T\varphi_k, \varphi_k \rangle|^2 \leqslant \sum_{j,\, k \in J} |\langle T\varphi_k, \varphi_j \rangle|^2. \end{aligned}$$

By Theorem 2.4.3 the last quantity is $||T||_2^2$; so (4) is satisfied, and equality occurs in (4) if and only if

$$(5) \qquad \langle T\varphi_k, \varphi_j\rangle = 0 \quad \text{unless } j = k \in K.$$

If equality occurs in (4), we deduce from (5) and Theorem 2.4.3. that

$$T = \sum_{k \in K} \langle T\varphi_k, \varphi_k\rangle (\varphi_k \otimes \varphi_k) = \sum_n \lambda_n(\psi_n \otimes \psi_n),$$

the series converging with respect to $||\cdot||_2$ and so also with respect to $||\cdot||$. Hence

$$(6) \qquad Tx = \sum_n \lambda_n \langle x, \psi_n\rangle \psi_n \quad (x \in \mathcal{H}),$$

and T is normal.

Suppose conversely that T is normal (and, of course, compact, since $T \in \mathcal{C}_2$). We can assume that the orthonormal sequence (ψ_n) was chosen, at the outset, so that (6) is satisfied, since this implies also that $\langle T\psi_n, \psi_n\rangle = \lambda_n$. It follows easily from (6) that (5) is satisfied and hence that equality occurs in (4).

The following result is an immediate consequence of Theorems 2.4.4 and 2.4.5. The first part is usually known as Schur's inequality, and the second part is due to Goldfain [23].

COROLLARY 2.4.6. *Suppose that $h \in L_2(E \times E, \mu \times \mu)$ and that T_h is the associated integral operator on $L_2(E, \mu)$ defined by*

$$(T_h f)(s) = \int_E h(s, t)f(t)d\mu(t).$$

Let (λ_n) be the sequence of non-zero eigenvalues of T_h, counted according to their algebraic multiplicities. Then

(i) $\sum_n |\lambda_n|^2 \leqslant \iint |h(s, t)|^2 d\mu(s)d\mu(t);$

(ii) equality occurs in (i) if and only if T_h is normal.

In view of Theorem 2.4.4, it is natural to ask for a simple description of the trace class operators on an L_2 space; but it seems that no such characterization exists. Suppose that $k(s, t)$ is continuous on the square $[0, 1] \times [0, 1]$, the associated integral operator K on $L_2(0, 1)$ (Lebesgue measure) is self-adjoint, (λ_n) is the sequence of non-zero eigenvalues of K (counted according to their multiplicities) and (φ_n) is the corresponding orthonormal sequence of eigenfunctions. It is not difficult to show that each φ_n is continuous [58: p. 22]. If K is a positive operator (in the sense of §1.7), a classical theorem of Mercer ([48]; see also [58: p. 128]) asserts that

$$k_n(s, t) = \sum_n \lambda_n \varphi_n(s)\overline{\varphi_n(t)},$$

the series being uniformly convergent for s, t in $[0, 1]$. Thus

$$\sum_n \lambda_n = \sum_n \int_0^1 \lambda_n |\varphi_n(s)|^2 ds = \int_0^1 k(s, s)ds < \infty.$$

It follows that K lies in the trace class (but this conclusion fails, in general, if we omit the assumption that K is positive).

CHAPTER 3

Fredholm Theory for Trace Class Operators

3.1 *Introduction*

The classical Fredholm theory is concerned with the solution of integral equations of the form

$$(1) \qquad f(s)-\lambda\int_a^b h(s,\ t)f(t)dt = g(s) \quad (a \leqslant s \leqslant b),$$

where λ is a complex number, $[a,\ b]$ is a compact real interval, and the known functions g, h and the unknown function f are all required to be continuous. For a full account of this subject, we refer to [43]; in this section we are concerned only to describe certain general features of the theory. The solution of equation (1) is expressed in terms of an entire function

$$(2) \qquad d(\lambda) = 1+\sum_1^\infty d_n\lambda^n$$

called the *Fredholm determinant* of the kernel h, and a kernel

$$(3) \qquad D_\lambda(s,\ t) = \sum_0^\infty D_n(s,\ t)\lambda^n,$$

the *first Fredholm minor*. It can be obtained, heuristically, by considering (1) as the limiting case of a finite system of linear equations in a finite number of unknowns; the functions $d(\lambda)$ and $D_\lambda(s,\ t)$ are the limits of certain determinants which arise when this

system is solved by Cramer's rule. The series in (3) converges for all s, t in $[a, b]$ and all complex λ, uniformly when λ is confined to a bounded subset of the complex plane. When $d(\lambda) \neq 0$, equation (1) has a unique solution given by

$$f(s) = g(s) + \frac{\lambda}{d(\lambda)} \int_a^b D_\lambda(s, t) g(t) \mathrm{d}t.$$

If $d(\lambda) = 0$ there is a non-trivial solution of the homogeneous equation

$$f(s) - \lambda \int_a^b h(s, t) f(t) dt = 0 \quad (a \leqslant s \leqslant b);$$

the conditions under which (1) has a solution, and the form of the general solution when it exists, can then be expressed in terms of Fredholm minors of higher orders.

There are explicit formulae for d_n and $D_n(s, t)$ in terms of h; for example,

$$d_1 = -\int_a^b h(s, s) ds. \tag{4}$$

By analogy with matrix theory the integral in (4) can be regarded as the 'trace' of the kernel h. It turns out that each of the coefficients d_n can be expressed in terms of such traces of h and its iterated kernels and that there is a similar expression, involving both the kernels and their traces, for D_n These formulae for d_n and D_n were given by Plemelj [50]. By use of closely analogous formulae we shall introduce the Fredholm determinant and the first Fredholm minor of a trace class operator in §3.3. We hasten to add that, if T_h is the integral operator in $L_2([a, b])$ associated with a continuous kernel h, then T_h need not be in the trace class.

If f, g and h are required only to be of class L_2, the theory just described requires modification; for example, the function $h(s, s)$ can fail to be integrable and equation (4) is then meaningless. Carleman [11] showed that, if h is redefined to be zero at all points of the form (s, s), then the classical Fredholm formulae remain valid; a simpler proof of this result was given by Smithies [57: see also 58]. Appropriately modified versions of Plemelj's formulae for d_n and D_n play an essential part in the method used by Smithies. In both the classical Fredholm theory and its modification for L_2 functions, it is possible to obtain estimates for the rates of growth of $d(\lambda)$ and D_λ as $|\lambda| \to \infty$.

Our purpose in this chapter is to provide a Fredholm theory for equations of the form $x - \lambda T x = y$, where T lies in the trace class of operators on a Hilbert space $\mathcal{H}$ and $x, y \in \mathcal{H}$. Such a theory, valid when T lies in an arbitrary von Neumann–Schatten class, has been expounded by Dunford and Schwartz [16 : XI 9]. Although the Fredholm formulae given in this chapter apply only to trace class operators, they can be used to obtain sharp estimates for the rate of growth on the resolvent of a quasi-nilpotent operator in a general von Neumann–Schatten class (see Theorem 3.4.6). Our method is a modification of the one used by Smithies [57, 58].

We note that a Fredholm theory for trace class operators on a Banach space has been developed by Ruston [53, 54], and that Grothendieck obtained similar results for trace class operators acting on locally convex spaces [24]. A sharp estimate for the rate of growth of the Fredholm determinant of a trace class operator acting on a Hilbert space was given by Lidskiĭ [39].

3.2. *Fredholm formulae for operators of finite rank*

In this section we develop a Fredholm theory for an operator of

finite rank on a Hilbert space $\mathcal{H}$. In §3.3 our results are extended, so as to apply to a trace class operator, by an approximation process.

We introduce some notation which will be used throughout §3.2. Let T be the operator defined by

$$(1) \qquad Tx = \sum_{j=1}^{n} \mu_j \langle x, \varphi_j \rangle \psi_j \quad (x \in \mathcal{H}),$$

where $(\varphi_1, \ldots, \varphi_n)$ and $(\psi_1, \ldots, \psi_n)$ are orthonormal systems in $\mathcal{H}$ and $\mu_1, \ldots, \mu_n$ are positive real numbers. The discussion at the end of §1.9 shows that each operator of finite rank can be represented in this way. We denote by $\mathfrak{M}$ the subspace generated by $\varphi_1, \ldots, \varphi_n$, $\psi_1, \ldots, \psi_n$. Since Tx is a linear combination of $\psi_1, \ldots, \psi_n$ for each x in $\mathcal{H}$, and $Ty = 0$ if y is orthogonal to each of $\varphi_1, \ldots, \varphi_n$, it follows that

$$(2) \qquad Tx \in \mathfrak{M} \quad (x \in \mathfrak{M}), \qquad Ty = 0 \quad (y \in \mathfrak{M}^{\perp});$$

we denote by $T^{\mathfrak{M}}$ the restriction of T to $\mathfrak{M}$. We extend $(\varphi_1, \ldots, \varphi_n)$ to an orthonormal basis $(\varphi_1, \ldots, \varphi_k)$ of $\mathfrak{M}$, and define a $k \times k$ matrix $A = [a_{ij}]$ by the condition

$$(3) \qquad T\varphi_j = \sum_{i=1}^{k} a_{ij}\varphi_i \quad (j = 1, \ldots, k).$$

Thus A is the matrix of the operator $T^{\mathfrak{M}}$ with respect to the basis $(\varphi_1, \ldots, \varphi_k)$. Since

$$||T\varphi_j||^2 = \sum_{i=1}^{k} |a_{ij}|^2$$

and, by (1), $T\varphi_j = \mu_j\psi_j$ or 0 according as $j \leqslant n$ or $j > n$, it follows that

$$(4) \qquad \sum_{i=1}^{k} |a_{ij}|^2 = \begin{cases} \mu_j^2 & (1 \leqslant j \leqslant n), \\ 0 & (n < j \leqslant k). \end{cases}$$

For each complex number λ we define

$$(5) \qquad d(\lambda) = \det (I_k - \lambda A),$$

the determinant of $I_k - \lambda A$, where I_k is the $k \times k$ identity matrix. The transposed cofactor matrix of $I_k - \lambda A$ will be denoted by $B(\lambda) = [b_{ij}(\lambda)]$, so that $b_{ij}(\lambda)$ is the (j, i) cofactor of $I_k - \lambda A$ and, by elementary matrix theory,

$$(6) \qquad (I_k - \lambda A)B(\lambda) = B(\lambda)(I_k - \lambda A) = d(\lambda)I_k.$$

We introduce a $k \times k$ matrix $C(\lambda) = [c_{ij}(\lambda)]$ and an operator D_λ on $\mathcal{H}$ by the equations

$$(7) \qquad C(\lambda) = AB(\lambda),$$

$$(8) \qquad D_\lambda x = \sum_{i,j=1}^{k} c_{ij}(\lambda) \langle x, \varphi_j \rangle \varphi_i \quad (x \in \mathcal{H}).$$

Then

$$D_\lambda \varphi_j = \sum_{i=1}^{k} c_{ij}(\lambda)\varphi_i,$$

and $D_\lambda y = 0$ whenever y is orthogonal to each of $\varphi_1, \ldots, \varphi_k$. It follows that

$$(9) \qquad D_\lambda x \in \mathfrak{M} \quad (x \in \mathfrak{M}), \qquad D_\lambda y = 0 \quad (y \in \mathfrak{M}^{\perp}),$$

and that the restriction $D_\lambda^{\mathfrak{M}}$ of D_λ to $\mathfrak{M}$ is the operator whose matrix with respect to the basis $\varphi_1, \ldots, \varphi_k$ is $C(\lambda)$.

Although $d(\lambda)$ and D_λ are polynomials in λ, the latter having operator-valued coefficients, it is more convenient for our purposes

to express them as power series by inserting additional terms with zero coefficients. We therefore write

$$(10) \qquad d(\lambda) = \sum_{m=0}^{\infty} \lambda^m d_m, \qquad D_\lambda = \sum_{m=0}^{\infty} \lambda^m D_m.$$

LEMMA 3.2.1. *For each complex number* λ

$$(11) \qquad (d(\lambda)I+\lambda D_\lambda)(I-\lambda T) = (I-\lambda T)(d(\lambda)I+\lambda D_\lambda) = d(\lambda)I.$$

Proof. It follows at once from (2) and (9) that each of the operators occurring in (11) leaves invariant the subspaces $\mathfrak{M}$ and $\mathfrak{M}^\perp$, and that

$$(d(\lambda)I+\lambda D_\lambda)(I-\lambda T)y = (I-\lambda T)(d(\lambda)I+\lambda D_\lambda)y = d(\lambda)y$$

when $y \in \mathfrak{M}^\perp$. It remains to prove the equations obtained from (11) by restricting each operator to $\mathfrak{M}$. For this purpose it is sufficient to derive the corresponding relations

$$(12) \qquad \begin{aligned}(d(\lambda)I_k+\lambda C(\lambda))(I_k-\lambda A)\\ = (I_k-\lambda A)(d(\lambda)I_k+\lambda C(\lambda)) = d(\lambda)I_k\end{aligned}$$

for the matrices of these (restricted) operators with respect to the basis $\varphi_1, \ldots, \varphi_k$. By (6)

$$d(\lambda)I_k+\lambda C(\lambda) = d(\lambda)I_k+\lambda AB(\lambda) = B(\lambda);$$

and this, together with another application of (6), yields (12).

The significance of Lemma 3.2.1 is that, when $d(\lambda) \neq 0$, $I-\lambda T$ has an inverse given by

$$(I-\lambda T)^{-1} = I+\lambda[d(\lambda)]^{-1}D_\lambda.$$

Our next objective is to obtain expressions for d_m and D_m, the coefficients in the expansions (10) of $d(\lambda)$ and D_λ, in a form which remains meaningful when T is replaced by a general trace class operator. For this purpose we need some further notation and an auxiliary result. Let

$$(13) \qquad \sigma_r = \tau(T^r) \quad (r = 1, 2, \ldots),$$

and suppose that the expression for the polynomial $d(\lambda) = \det(I_k - \lambda A)$ as a product of linear factors is

$$(14) \qquad d(\lambda) = (1-\lambda\theta_1)(1-\lambda\theta_2)\ldots(1-\lambda\theta_k).$$

Thus $\theta_1, \ldots, \theta_k$ are the eigenvalues of the matrix A, counted according to their algebraic multiplicities.

LEMMA 3.2.2. $\sigma_r = \theta_1^r+\theta_2^r+\ldots+\theta_k^r$ $(r = 1, 2, \ldots)$.

Proof. Suppose that S is any operator on $\mathcal{H}$ which leaves $\mathfrak{M}$ invariant and annihilates $\mathfrak{M}^\perp$; and let $P = [p_{ij}]$ be the matrix, with respect to the basis $\varphi_1, \ldots, \varphi_k$, of the operator $S^{\mathfrak{M}}$ obtained by restricting S to $\mathfrak{M}$. Let $\{\xi_j : j \in J\}$ be an orthonormal basis of $\mathcal{H}$ which contains each of $\varphi_1, \ldots, \varphi_k$. Then S has finite rank and so lies in the trace class, $S\xi_j = 0$ except when ξ_j is one of $\varphi_1, \ldots, \varphi_k$, and therefore

$$\begin{aligned}\tau(S) &= \sum_{j\in J} \langle S\xi_j, \xi_j\rangle\\ &= \sum_{i=1}^{k} \langle S\varphi_i, \varphi_i\rangle = \sum_{i=1}^{k} p_{ii}.\end{aligned}$$

It follows from elementary matrix theory that $\tau(S)$ is the sum of the eigenvalues of P. If we take $S = T^r$ then $S^{\mathfrak{M}}$ is $(T^{\mathfrak{M}})^r$ and its matrix P is A^r. The eigenvalues of P, counted according to their algebraic multiplicities, are $\theta_1^r, \theta_2^r, \ldots, \theta_k^r$, and so

$$\tau(T^r) = \theta_1^r+\theta_2^r+\ldots+\theta_k^r.$$

LEMMA 3.2.3. *The coefficients in the expansion* (10) *of* $d(\lambda)$ *are given by* $d_0 = 1$ *and*

$$(15)\qquad d_m = \frac{(-1)^m}{m!}\begin{vmatrix} \sigma_1 & 1 & 0 & \ldots & 0 & 0 \\ \sigma_2 & \sigma_1 & 2 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \sigma_{m-1} & \sigma_{m-2} & \sigma_{m-3} & \ldots & \sigma_1 & m-1 \\ \sigma_m & \sigma_{m-1} & \sigma_{m-2} & \ldots & \sigma_2 & \sigma_1 \end{vmatrix}\quad (m \geqslant 1).$$

Proof. By differentiating (14) with respect to λ we obtain

$$\frac{d'(\lambda)}{d(\lambda)} = -\sum_{j=1}^{k} \frac{\theta_j}{1-\lambda\theta_j} = -\sum_{j=1}^{k}\sum_{m=0}^{\infty} \lambda^m\theta_j^{m+1} = -\sum_{m=0}^{\infty}\lambda^m\sigma_{m+1},$$

this expansion being valid for sufficiently small λ. Since

$$d(\lambda) = \sum_{m=0}^{\infty}\lambda^m d_m, \qquad d'(\lambda) = \sum_{m=0}^{\infty} m\lambda^{m-1}d_m,$$

we have

$$\sum_{m=1}^{\infty} m\lambda^{m-1}d_m = -\left(\sum_{m=0}^{\infty}\lambda^m d_m\right)\left(\sum_{m=0}^{\infty}\lambda^m\sigma_{m+1}\right).$$

By equating coefficients of λ^{r-1} we obtain

$$rd_r = -(d_0\sigma_r+d_1\sigma_{r-1}+\ldots+d_{r-1}\sigma_1).$$

Hence

$$(16)\qquad \begin{cases} d_1 & = -d_0\sigma_1, \\ d_1\sigma_1 + 2d_2 & = -d_0\sigma_2, \\ d_1\sigma_2 + d_2\sigma_1 + 3d_3 & = -d_0\sigma_3, \\ \ldots & \ldots \\ d_1\sigma_{m-1}+d_2\sigma_{m-2}+d_3\sigma_{m-3}+\ldots+d_{m-1}\sigma_1+md_m & = -d_0\sigma_m. \end{cases}$$

It follows from (5) that $d_0 = d(0) = 1$. By applying Cramer's rule to the system (16) of linear equations in $d_1, \ldots, d_m$ we obtain (15).

LEMMA 3.2.4. *The coefficients in the power series* (10) *for* D_λ *are given by* $D_0 = T$ *and*

$$(17)\qquad D_m = \frac{(-1)^m}{m!}\begin{vmatrix} \sigma_1 & 1 & 0 & \ldots & 0 & 0 \\ \sigma_2 & \sigma_1 & 2 & \ldots & 0 & 0 \\ \sigma_3 & \sigma_2 & \sigma_1 & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ \sigma_m & \sigma_{m-1} & \sigma_{m-2} & \ldots & \sigma_1 & m \\ T^{m+1} & T^m & T^{m-1} & \ldots & T^2 & T \end{vmatrix}\quad (m \geqslant 1),$$

where the determinant is to be interpreted as the operator obtained by formal expansion.

Proof. We begin by remarking that, because of the uniqueness assertion in Theorem 1.4.1, the process of 'equating coefficients' is valid for power series with operator-valued coefficients (in proving the present lemma we shall in fact be dealing with polynomials). It follows from (11) that

$$D_\lambda = \lambda TD_\lambda+d(\lambda)T,$$

and by equating coefficients of powers of λ we deduce that

(18) $$D_0 = T, \qquad D_m = TD_{m-1}+d_mT \quad (m \geqslant 1).$$

The recurrence relation (18) determines D_m for all $m = 0, 1, 2, \ldots$ Let E_m denote the right-hand side of (17) when $m \geqslant 1$, and define $E_0 = T$. If we expand the determinant in (17) by its last column we find that

$$E_m = TE_{m-1}+d_mT \quad (m \geqslant 1).$$

Hence E_m, as well as D_m, satisfies (18); and therefore $D_m = E_m$ $(m = 0, 1, 2, \ldots)$.

In §3.3 we shall use equations (13), (15), (17) and (10), in that order, to *define* the Fredholm determinant $d(\lambda)$ and the first Fredholm minor D_λ for a general trace class operator T. In that context, $d(\lambda)$ D_λ need not be polynomials; and, in order to show that they are entire functions, we require estimates for the orders of magnitude of the coefficients d_m and D_m. We shall also be interested in the rates of growth of these functions as $|\lambda| \to \infty$. For the present we continue studying the case in which T has finite rank, and in Lemmas 3.2.6, 3.2.7 and 3.2.8 we obtain the appropriate estimates in this case.

LEMMA 3.2.5. *(Hadamard's inequality). Let $P = [p_{ij}]$ be a $k \times k$ complex matrix. Then*

$$|\det P|^2 \leqslant \prod_{j=1}^{k} \left(\sum_{i=1}^{k} |p_{ij}|^2\right).$$

Proof. The result is apparent if $\det P = 0$, so we assume henceforth that P is non-singular. The 'columns' $v_j = (p_{1j}, p_{2j}, \ldots, p_{kj})$ $(j = 1, \ldots, k)$ of P form a basis in the Hilbert space C^k of k-tuples of complex numbers. By applying the Schmidt orthogonalization process to $v_1, \ldots, v_k$ we obtain an orthonormal basis $w_1, \ldots, w_k$ of C^k with the property that

(19) $$v_j = \sum_{r=1}^{j} t_{rj}w_r \quad (j = 1, \ldots, k)$$

for suitable scalars t_{rj} $(1 \leqslant r \leqslant j \leqslant k)$. If w_r is the vector $(u_{1r}, u_{2r}, \ldots, u_{kr})$ and we define $t_{rj} = 0$ when $r > j$, then (19) asserts that

$$p_{ij} = \sum_{r=1}^{k} u_{ir}t_{rj} \quad (i, j = 1, \ldots, k).$$

Hence $P = UT$, where U is the unitary $k \times k$ matrix $[u_{ij}]$, and T is the upper-triangular $k \times k$ matrix $[t_{ij}]$. Since $T = U^*P$ we have

$$\begin{aligned} |t_{jj}|^2 &= \left|\sum_{i=1}^{k} \bar{u}_{ij}p_{ij}\right|^2 \\ &\leqslant \left(\sum_{i=1}^{k} |u_{ij}|^2\right)\left(\sum_{i=1}^{k} |p_{ij}|^2\right) \\ &= \sum_{i=1}^{k} |p_{ij}|^2. \end{aligned}$$

Hence

$$\begin{aligned} |\det P|^2 &= |\det U|^2|\det T|^2 \\ &= |\det T|^2 \\ &= \prod_{j=1}^{k} |t_{jj}|^2 \\ &\leqslant \prod_{j=1}^{k} \left(\sum_{i=1}^{k} |p_{ij}|^2\right). \end{aligned}$$

In the next few lemmas, the positive integer n and the positive real numbers $\mu_1, \ldots, \mu_n$ are those occurring in (1), and Γ_p is defined by

(20) $$\Gamma_p = \sup_{s>0} s^{-p}\log(1+s) \quad (0 < p < 1).$$

Since $s^{-p}\log(1+s)$ is continuous on the open interval $(0, \infty)$ and has limit zero as s tends to either endpoint, Γ_p is finite.

Furthermore

$$1+s \leqslant \exp(\Gamma_p s^p) \quad (s > 0;\ 0 < p < 1). \tag{21}$$

LEMMA 3.2.6. *Suppose that r is an integer such that $1 \leqslant r < n$, and $0 < p < 1$. Then*

$$|d(\lambda)| \leqslant \prod_{j=1}^{n} (1+|\lambda|\mu_j), \tag{22}$$

$$|d(\lambda)| \leqslant \exp(|\lambda| \sum_{j=r+1}^{n} \mu_j) \prod_{j=1}^{r} (1+|\lambda|\mu_j), \tag{23}$$

$$|d(\lambda)| \leqslant \exp(|\lambda| \sum_{j=1}^{n} \mu_j), \tag{24}$$

$$|d(\lambda)| \leqslant \exp(\Gamma_p|\lambda|^p \sum_{j=r+1}^{n} \mu_j^p) \prod_{j=1}^{r} (1+|\lambda|\mu_j), \tag{25}$$

$$|d(\lambda)| \leqslant \exp(\Gamma_p|\lambda|^p \sum_{j=1}^{n} \mu_j^p). \tag{26}$$

Proof. Since $d(\lambda) = \det(I_k-\lambda A)$, Hadamard's inequality gives

$$\begin{aligned}|d(\lambda)|^2 &\leqslant \prod_{j=1}^{k} \{|\lambda|^2|a_{1j}|^2+|\lambda|^2|a_{2j}|^2+\dots \\ &\qquad +|1-\lambda a_{jj}|^2+\dots+|\lambda|^2|a_{kj}|^2\} \\ &\leqslant \prod_{j=1}^{k} \{1+2|\lambda|\ |a_{jj}|+|\lambda|^2 \sum_{i=1}^{k} |a_{ij}|^2\}.\end{aligned}$$

It follows from (4) that

$$\begin{aligned}|d(\lambda)|^2 &\leqslant \prod_{j=1}^{n} \{1+2|\lambda|\mu_j+|\lambda|^2\mu_j^2\} \\ &= \prod_{j=1}^{n} (1+|\lambda|\mu_j)^2.\end{aligned}$$

This proves (22). Since $1+s \leqslant \exp(s)$ whenever $s > 0$, the factor $1+|\lambda|\mu_j$ in the right-hand side of (22) can be replaced by $\exp(|\lambda|\mu_j)$; by making this change for appropriate values of j we obtain (23) and (24). In the same way (25) and (26) can be deduced from (22) by use of the inequality (21).

LEMMA 3.2.7. *Suppose that r is an integer such that $1 \leqslant r < n$ and $0 < p < 1$. Then*

$$||d(\lambda)I+\lambda D_\lambda|| \leqslant \exp(\tfrac12) \prod_{j=1}^{n} (1+|\lambda|\mu_j), \tag{27}$$

$$||d(\lambda)I+\lambda D_\lambda|| \leqslant \exp(\tfrac12+|\lambda| \sum_{j=r+1}^{n} \mu_j) \prod_{j=1}^{r} (1+|\lambda|\mu_j), \tag{28}$$

$$||d(\lambda)I+\lambda D_\lambda|| \leqslant \exp(\tfrac12+|\lambda| \sum_{j=1}^{n} \mu_j), \tag{29}$$

$$||d(\lambda)I+\lambda D_\lambda|| \leqslant \exp(\tfrac12+\Gamma_p|\lambda|^p \sum_{j=r+1}^{n} \mu_j^p) \prod_{j=1}^{r} (1+|\lambda|\mu_j), \tag{30}$$

$$||d(\lambda)I+\lambda D_\lambda|| \leqslant \exp(\tfrac12+\Gamma_p|\lambda|^p \sum_{j=1}^{n} \mu_j^p). \tag{31}$$

Proof. It is sufficient to prove (27), since the other results follow from (27) in the same way that the last four inequalities in Lemma 3.2.6 were deduced from (22). By (9), $d(\lambda)I+\lambda D_\lambda$ leaves invariant both the subspaces $\mathfrak{M}$ and $\mathfrak{M}^\perp$; so, in order to prove (27), it suffices to establish the corresponding inequalities for the operators obtained by restricting $d(\lambda)I+\lambda D_\lambda$ to these subspaces. Since $D_\lambda y = 0$ when $y \in \mathfrak{M}^\perp$, the restriction to $\mathfrak{M}^\perp$ is $d(\lambda)I$, and the estimate (22) for its norm $|d(\lambda)|$ is in fact better than we require. It remains to consider the restriction $d(\lambda)I+\lambda D_\lambda^{\mathfrak{M}}$ to $\mathfrak{M}$.

The matrix of $d(\lambda)I+\lambda D_\lambda^{\mathfrak{M}}$ with respect to the basis $\varphi_1, \dots, \varphi_k$ of $\mathfrak{M}$ is $d(\lambda)I_k+\lambda C(\lambda)$, and we have seen in proving Lemma 3.2.1 that this is $B(\lambda) = [b_{ij}(\lambda)]$, the transposed cofactor matrix of $I_k-\lambda A$. *Thus*

$$[d(\lambda)I+\lambda D_\lambda^{\mathfrak{M}}]\varphi_j = \sum_{i=1}^{k} b_{ij}(\lambda)\varphi_i,$$

$$\langle [d(\lambda)I+\lambda D_\lambda^{\mathfrak{M}}]\varphi_j,\ \varphi_i\rangle = b_{ij}(\lambda).$$

Given any unit vectors x and y in $\mathfrak{M}$, we may set

$$x = \alpha_1\varphi_1+\ldots+\alpha_k\varphi_k, \qquad y = \beta_1\varphi_1+\ldots+\beta_k\varphi_k,$$

where

$$|\alpha_1|^2+\ldots+|\alpha_k|^2 = |\beta_1|^2+\ldots+|\beta_k|^2 = 1.$$

Then

$$\langle [d(\lambda)I+\lambda D_\lambda^{\mathfrak{M}}]x,\ y\rangle = \sum_{i,j=1}^{k} b_{ij}(\lambda)\alpha_j\bar{\beta}_i$$

$$= -\begin{vmatrix} 0 & \bar{\beta}_1 & \cdots & \bar{\beta}_k \\ \alpha_1 & & & \\ \vdots & & I_k-\lambda A & \\ \alpha_k & & & \end{vmatrix}$$

(this last determinant reduces to the summation in the preceding line when it is expanded by the first row and the resulting determinants are then expanded by the first column). By Hadamard's inequality and equation (4)

$$|\langle [d(\lambda)I+\lambda D_\lambda^{\mathfrak{M}}]x,\ y\rangle|^2$$

$$\leqslant (\sum_{i=1}^{k}|\alpha_i|^2)\prod_{j=1}^{k}\{|\beta_j|^2+|\lambda|^2|a_{1j}|^2+$$

$$+|\lambda|^2|a_{2j}|^2+\ldots+|1-\lambda a_{jj}|^2+\ldots+|\lambda|^2|a_{kj}|^2\}$$

$$\leqslant \prod_{j=1}^{k}\{1+2|\lambda|\ |a_{jj}|+|\lambda|^2\sum_{i=1}^{k}|a_{ij}|^2+|\beta_j|^2\}$$

$$\leqslant \prod_{j=1}^{k}\{1+2|\lambda|\ |a_{jj}|+|\lambda|^2\sum_{i=1}^{k}|a_{ij}|^2\}\{1+|\beta_j|^2\}$$

$$\leqslant \prod_{j=1}^{k}\{1+2|\lambda|\mu_j+|\lambda|^2\mu_j^2\}\prod_{j=1}^{k}\exp(|\beta_j|^2)$$

$$= e\prod_{j=1}^{n}(1+|\lambda|\mu_j)^2.$$

Thus

$$|\langle [d(\lambda)I+\lambda D_\lambda^{\mathfrak{M}}]x,\ y\rangle| \leqslant \exp(\tfrac{1}{2})\prod_{j=1}^{n}(1+|\lambda|\mu_j)$$

whenever x and y are unit vectors in $\mathfrak{M}$. It follows that

$$||d(\lambda)I+\lambda D_\lambda^{\mathfrak{M}}|| \leqslant \exp(\tfrac{1}{2})\prod_{j=1}^{n}(1+|\lambda|\mu_j).$$

This completes the proof of (27).

LEMMA 3.2.8. *For* $m = 1, 2, 3, \ldots$

(32) $$|d_m| \leqslant m^{-m}\exp(m)\,(\sum_{j=1}^{n}\mu_j)^m,$$

(33) $$||d_mI+D_{m-1}|| \leqslant m^{-m}\exp(m+\tfrac{1}{2})\,(\sum_{j=1}^{n}\mu_j)^m.$$

Proof. We shall use the following result [60: p. 84]: if

$$f(\lambda) = \sum_{m=0}^{\infty} a_m\lambda^m$$

is an entire function, $r > 0$ and $M(r) = \sup\{|f(\lambda)| : |\lambda| = r\}$, then $|a_m| \leqslant r^{-m}M(r)$ $(m = 0, 1, 2, \ldots)$.

It follows from (24) that, for the entire function $d(\lambda)$,

$$M(r) \leqslant \exp(r\sum_{j=1}^{n}\mu_j).$$

Thus

$$|d_m| \leqslant r^{-m}\exp(r\sum_{j=1}^{n}\mu_j),$$

and by taking $r = m(\sum_{j=1}^{n} \mu_j)^{-1}$ we obtain (32).

Suppose next that x and y are unit vectors in $\mathcal{H}$ and let

$$\begin{aligned} f(\lambda) &= \langle [d(\lambda)I+\lambda D_\lambda]x,\ y\rangle \\ &= \langle x,\ y\rangle + \sum_{m=1}^{\infty} \lambda^m \langle [d_m I + D_{m-1}]x,\ y\rangle. \end{aligned}$$

It follows from (29) that, for this function,

$$M(r) \leqslant \exp(\tfrac{1}{2}+r \sum_{j=1}^{n} \mu_j).$$

Hence

$$|\langle [d_m I+D_{m-1}]x,\ y\rangle| \leqslant r^{-m} \exp(\tfrac{1}{2}+r \sum_{j=1}^{n} \mu_j);$$

and since the last inequality has been proved for arbitrary unit vectors x and y,

$$||d_m I+D_{m-1}|| \leqslant r^{-m} \exp(\tfrac{1}{2}+r \sum_{j=1}^{n} \mu_j).$$

This reduces to (33) when $r = m(\sum_{j=1}^{n} \mu_j)^{-1}$.

3.3. *Fredholm formulae for trace class operators*

In this section we define the Fredholm determinant and the first Fredholm minor of a trace class operator acting on a Hilbert space, and show that much of the theory developed in §3.2 remains valid in this more general setting. The main results are Theorems 3.3.9, 3.3.10 and 3.3.13.

We introduce some notation that will be used throughout §3.3. Let T be a trace class operator acting on a Hilbert space $\mathcal{H}$. It follows from Theorem 1.9.3 that there exist (finite or infinite) orthonormal sequences (φ_j) and (ψ_j) in $\mathcal{H}$, and a sequence (μ_j) of

positive real numbers (the eigenvalues of $(T^*T)^{1/2}$) such that

(1) $$Tx = \sum_j \mu_j \langle x,\ \varphi_j\rangle \psi_j \quad (x \in \mathcal{H}).$$

By equation 2.3(4),

(2) $$\sum_j \mu_j = ||T||_1.$$

We shall usually assume that the index j runs through the set of all positive integers; for if the summation in (1) is finite, then T has finite rank and our results are already contained in those of §3.2.

Since the trace class $\mathcal{C}_1$ is an ideal in $\mathcal{B}(\mathcal{H})$, $T^m \in \mathcal{C}_1$ $(m = 1, 2, \ldots)$. We define

(3) $$\sigma_m(T) = \tau(T^m) \quad (m = 1, 2, \ldots),$$

(4) $$d_m(T) = \frac{(-1)^m}{m!} \begin{vmatrix} \sigma_1 & 1 & 0 & \cdots & 0 & 0 \\ \sigma_2 & \sigma_1 & 2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_{m-1} & \sigma_{m-2} & \sigma_{m-3} & \cdots & \sigma_1 & m-1 \\ \sigma_m & \sigma_{m-1} & \sigma_{m-2} & \cdots & \sigma_2 & \sigma_1 \end{vmatrix}$$

$(m = 1, 2, \ldots)$ and $d_0(T) = 1$,

(5) $$D_m(T) = \frac{(-1)^m}{m!} \begin{vmatrix} \sigma_1 & 1 & 0 & \cdots & 0 & 0 \\ \sigma_2 & \sigma_1 & 2 & \cdots & 0 & 0 \\ \sigma_3 & \sigma_2 & \sigma_1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \sigma_m & \sigma_{m-1} & \sigma_{m-2} & \cdots & \sigma_1 & m \\ T^{m+1} & T^m & T^{m-1} & \cdots & T^2 & T \end{vmatrix}$$

$(m = 1, 2, \ldots)$ and $D_0(T) = T$, where $\sigma_m = \sigma_m(T)$ in (4) and (5).

We shall consider T as the limit of the operators $T_1, T_2, \ldots,$ of finite rank, which are defined by

$$T_n x = \sum_{j=1}^{n} \mu_j \langle x, \varphi_j \rangle \psi_j \quad (x \in \mathcal{H}). \tag{6}$$

Corresponding to equation (2), we have

$$\sum_{j=1}^{n} \mu_j = ||T_n||_1 \leqslant ||T||_1. \tag{7}$$

LEMMA 3.3.1. *For* $m = 1, 2, 3, \ldots,$

(i) $||T^m - T_n^m||_1 \to 0$ as $n \to \infty$,

(ii) $\sigma_m(T_n) \to \sigma_m(T)$ as $n \to \infty$.

For $m = 0, 1, 2, \ldots,$

(iii) $d_m(T_n) \to d_m(T)$ as $n \to \infty$,

(iv) $||D_m(T) - D_m(T_n)||_1 \to 0$ as $n \to \infty$.

Proof. In proving the final assertion of Theorem 2.3.8 (that is dense in $\mathcal{C}_p$) we have already shown that

$$||T - T_n||_1 \to 0 \text{ as } n \to \infty,$$

whence $||T_n||_1 \to ||T||_1$. Now

$$\begin{aligned} ||T^m - T_n^m||_1 &= ||\sum_{r=1}^{m} T_n^{r-1}(T-T_n)T^{m-r}||_1 \\ &\leqslant \sum_{r=1}^{m} ||T_n^{r-1}(T-T_n)T^{m-r}||_1. \end{aligned}$$

It follows from Corollary 2.3.11 and the inequality 2.3(5) that

$$\begin{aligned} ||T^m - T_n^m||_1 &\leqslant \sum_{r=1}^{m} ||T_n||^{r-1} ||T-T_n||_1 ||T||^{m-r} \\ &\leqslant \sum_{r=1}^{m} ||T_n||_1^{r-1} ||T-T_n||_1 ||T||_1^{m-r} \\ &\to 0 \text{ as } n \to \infty. \end{aligned}$$

This proves (i), and (ii) follows immediately since, by Lemma 2.3.3,

$$|\sigma_m(T) - \sigma_m(T_n)| = |\tau(T^m - T_n^m)| \leqslant ||T^m - T_n^m||_1.$$

The last two parts of the lemma are obvious consequences of (i), (ii) and the definitions (4) and (5) of $d_m(T)$ and $D_m(T)$.

LEMMA 3.3.2. *For* $m = 1, 2, 3, \ldots,$

(i) $|d_m(T)| \leqslant m^{-m} \exp(m) \, ||T||_1^m;$

(ii) $||D_m(T)||_1 \leqslant m^{-m} \exp(m+\tfrac{1}{2}) \, ||T||_1^{m+1}.$

Proof. By Lemma 3.2.8 and the inequality (7)

$$\begin{aligned} |d_m(T_n)| &\leqslant m^{-m} \exp(m) \left(\sum_{j=1}^{n} \mu_j\right)^m \\ &\leqslant m^{-m} \exp(m) \, ||T||_1^m. \end{aligned}$$

By letting $n \to \infty$ and using Lemma 3.3.1(iii), we obtain (i).

By equation 3.2(18), Lemma 3.2.8. and equation (2) we have

$$\begin{aligned} ||D_m(T_n)||_1 &= ||T_n(D_{m-1}(T_n) + d_m(T_n)I)||_1 \\ &\leqslant ||T_n||_1 ||D_{m-1}(T_n) + d_m(T_n)I|| \\ &\leqslant ||T_n||_1 \, m^{-m} \exp(m+\tfrac{1}{2}) \left(\sum_{j=1}^{n} \mu_j\right)^m \\ &\leqslant ||T_n||_1 \, m^{-m} \exp(m+\tfrac{1}{2}) \, ||T||_1^m; \end{aligned}$$

and (ii) follows, from Lemma 3.3.1(i) and (iv), by letting $n \to \infty$.

We deduce from Lemma 3.3.2 that the series

$$\sum_{m=0}^{\infty} \lambda^m d_m(T), \quad \sum_{m=0}^{\infty} |\lambda|^m \, ||D_m(T)||_1$$

converge for all complex λ. The entire function

$$d(\lambda, T) = \sum_{m=0}^{\infty} \lambda^m d_m(T) \tag{8}$$

will be called the *Fredholm determinant* of T. Since $\mathcal{C}_1$ is complete the series $\Sigma\, \lambda^m D_m(T)$ converges with respect to $||\cdot||_1$ for all complex λ. The trace class operator $D_{\lambda, T}$ defined by

$$D_{\lambda, T} = \sum_{m=0}^{\infty} \lambda^m D_m(T) \tag{9}$$

will be called the *first Fredholm minor* of T.

LEMMA 3.3.3. *(i)* $d(\lambda, T_n) \to d(\lambda, T)$ *as* $n \to \infty$.

(ii) $\quad ||D_{\lambda,T}-D_{\lambda,T_n}||_1 \to 0$ *as* $n \to \infty$.

(iii) $\quad ||D_{\lambda,T}-D_{\lambda,T_n}|| \to 0$ *as* $n \to \infty$.

Proof. We shall prove only part (ii) of this lemma, since (i) follows from a very similar argument and (iii) is an immediate consequence of (ii).

For each positive integer q,

$$\begin{aligned} ||D_{\lambda,T}-D_{\lambda,T_n}||_1 &= ||\sum_{m=0}^{\infty} \lambda^m[D_m(T)-D_m(T_n)]||_1 \\ &\leqslant \sum_{m=0}^{q} |\lambda|^m ||D_m(T)-D_m(T_n)||_1 + \\ &\quad + \sum_{m=q+1}^{\infty} |\lambda|^m[||D_m(T)||_1+||D_m(T_n)||_1]. \end{aligned}$$

By Lemma 3.3.2(ii), applies to both T and T_n, together with the inequality (7), we have

$$\begin{aligned} ||D_{\lambda,T}-D_{\lambda,T_n}||_1 &\leqslant \sum_{m=0}^{q} |\lambda|^m ||D_m(T)-D_m(T_n)||_1 + \\ &\quad +2 \sum_{m=q+1}^{\infty} |\lambda|^m m^{-m} \exp(m+½)||T||_1^{m+1}. \end{aligned} \tag{10}$$

Given a positive ϵ, we can choose q so that

$$2 \sum_{m=q+1}^{\infty} |\lambda|^m m^{-m} \exp(m+½)||T||_1^{m+1} < ½\epsilon.$$

By Lemma 3.3.1(iv) there is a positive integer n_0 such that

$$\sum_{m=0}^{q} |\lambda|^m ||D_m(T)-D_m(T_n)||_1 < ½\epsilon$$

whenever $n \geqslant n_0$. It follows from (10) that

$$||D_{\lambda,T}-D_{\lambda,T_n}||_1 < \epsilon$$

whenever $n \geqslant n_0$.

LEMMA 3.3.4. *For all complex* λ,

$$\begin{aligned} &[d(\lambda,T)I+\lambda D_{\lambda,T}](I-\lambda T) \\ &\quad = (I-\lambda T)[d(\lambda,T)I+\lambda D_{\lambda,T}] = d(\lambda,T)I. \end{aligned} \tag{11}$$

Proof. By virtue of Lemmas 3.3.3(i), (iii) and 3.3.1(i), this result follows at once from the corresponding relations for T_n (see Lemma 3.2.1).

COROLLARY 3.3.5. *If* $d(\lambda, T) \neq 0$ *then* $I-\lambda T$ *has an inverse given by*

$$(I-\lambda T)^{-1} = I+\lambda[d(\lambda, T)]^{-1}D_{\lambda,T}. \tag{12}$$

Our next aim is to show that, if $d(\lambda_0, T) = 0$ and λ_0 has multiplicity k as a zero of $d(\lambda, T)$, then λ_0^{-1} is an eigenvalue of T and has algebraic multiplicity k. Before proving this we require two auxiliary results. In the following lemma, $d'(\lambda, T)$ denotes the derivative, with respect to λ, of $d(\lambda, T)$.

LEMMA 3.3.6. $\tau(D_{\lambda,T}) = -d'(\lambda, T)$.

Proof. It follows from (5), (4) and the comments following those equations that

$$\tau(D_0(T)) = \tau(T) = \sigma_1 = -d_1(T),$$

$$\tau(D_m(T)) = \frac{(-1)^m}{m!}\begin{vmatrix} \sigma_1 & 1 & 0 & \dots & 0 & 0 \\ \sigma_2 & \sigma_1 & 2 & \dots & 0 & 0 \\ \sigma_3 & \sigma_2 & \sigma_1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \sigma_m & \sigma_{m-1} & \sigma_{m-2} & \dots & \sigma_1 & m \\ \sigma_{m+1} & \sigma_m & \sigma_{m-1} & \dots & \sigma_2 & \sigma_1 \end{vmatrix}$$

$$= -(m+1)\,d_{m+1}(T) \quad (m = 1, 2, \dots).$$

Since the series (9) for $D_{\lambda,T}$ converges with respect to $||\cdot||_1$ and τ is a continuous linear functional on $\mathcal{C}_1$, we have

$$\tau(D_{\lambda,T}) = \sum_{m=0}^{\infty} \lambda^m \tau(D_m(T))$$

$$= -\sum_{m=0}^{\infty} (m+1)\lambda^m d_{m+1}(T)$$

$$= -d'(\lambda, T).$$

LEMMA 3.3.7. *Let K be a trace class operator acting on a Hilbert space $\mathcal{H}$ and let $\mathfrak{D}$ be the open set consisting of all non-zero complex numbers λ for which λ^{-1} is not an eigenvalue of K. For each λ in $\mathfrak{D}$ define an operator K_λ on $\mathcal{H}$ by*

$$(I-\lambda K)^{-1} = I+\lambda K_\lambda.$$

Then K_λ is a trace class operator; and if $\lambda_0 \in \mathfrak{D}$ then $\tau(K_\lambda)$ is bounded on some neighbourhood of λ_0.

Proof. Since $(I-\lambda K)(I+\lambda K_\lambda) = I$ we have

$$K_\lambda = K(I+\lambda K_\lambda);$$

so K_λ is in the trace class and

(13) $$|\tau(K_\lambda)| \leqslant ||K(I+\lambda K_\lambda)||_1 \leqslant ||K||_1||I+\lambda K_\lambda||.$$

By Theorem 1.5.1, the operator-valued function $\lambda \to I+\lambda K_\lambda = (I-\lambda K)^{-1}$ is analytic on $\mathfrak{D}$, and is therefore continuous with respect to the norm topology on $\mathcal{B}(\mathcal{H})$. It follows that, if $\lambda_0 \in \mathfrak{D}$, then $||I+\lambda K_\lambda||$ is bounded on some neighbourhood of λ_0; and, by (13), the same is true of $\tau(K_\lambda)$.

LEMMA 3.3.8. *Let λ_0 be a complex number such that $d(\lambda_0, T)=0$. Then λ_0^{-1} is an eigenvalue of T and its algebraic multiplicity as such is equal to the multiplicity of λ_0 as a zero of $d(\lambda, T)$.*

Proof. Suppose that λ_0 has multiplicity k as a zero of $d(\lambda, T)$. Then

(14) $$d(\lambda, T) = (\lambda-\lambda_0)^k g(\lambda),$$

where g is an entire function and $g(\lambda_0) \neq 0$. We may choose a positive δ such that $g(\lambda) \neq 0$ when $|\lambda-\lambda_0| \leqslant \delta$. If $0 < |\lambda-\lambda_0| \leqslant \delta$ then $d(\lambda, T) \neq 0$ and, by Corollary 3.3.5,

(15) $$(I-\lambda T)^{-1} = I+\lambda S_\lambda \quad (0 < |\lambda-\lambda_0| \leqslant \delta),$$

where

(16) $$S_\lambda = [d(\lambda, T)]^{-1}D_{\lambda,T} \quad (0 < |\lambda-\lambda_0| \leqslant \delta).$$

By Lemma 3.3.6, $\tau(S_\lambda) = -d'(\lambda, T)/d(\lambda, T)$; and from (14) we have

(17) $$\tau(S_\lambda) = \frac{k}{\lambda_0-\lambda} - h(\lambda) \quad (0 < |\lambda-\lambda_0| \leqslant \delta),$$

where $h(\lambda) = g'(\lambda)/g(\lambda)$ and thus h is bounded near λ_0.

Since $\tau(S_\lambda)$ is unbounded on each neighbourhood of λ_0 and $\lambda_0 \neq 0$ (because $d(0, T) = 1$), it follows from Lemma 3.3.7 that λ_0^{-1} is an

eigenvalue of T. Suppose that, as such, it has algebraic multiplicity m. By Theorem 1.8.1 there exist closed subspaces $\mathfrak{N}$ and $\mathfrak{R}$ of $\mathfrak{H}$, both invariant under T, such that

$$\mathfrak{N}+\mathfrak{R} = \mathfrak{H}, \qquad \mathfrak{N}\cap\mathfrak{R} = (0), \tag{18}$$

$\mathfrak{N}$ is m-dimensional; and, if $T^{\mathfrak{N}}$ and $T^{\mathfrak{R}}$ denote the restrictions of T to $\mathfrak{N}$ and $\mathfrak{R}$, respectively, then $I-\lambda_0 T^{\mathfrak{N}}$ is nilpotent and $I-\lambda_0 T^{\mathfrak{R}}$ is invertible. Since an orthonormal system in $\mathfrak{R}$ is also an orthonormal system in $\mathfrak{H}$, it follows at once from Definition 2.1.1 that $T^{\mathfrak{R}}$ is in the trace class of operators on $\mathfrak{R}$. We shall use the decomposition of T relative to $\mathfrak{N}$ and $\mathfrak{R}$ to obtain a second expression for $\tau(S_\lambda)$ which, when compared with (17), yields the required equation $k = m$.

By Theorem 1.8.3(iii), any closed subspace of $\mathfrak{H}$ which is invariant under T is invariant also under $(I-\lambda T)^{-1}$ for all complex numbers λ such that this inverse operator exists. In particular, $\mathfrak{N}$ and $\mathfrak{R}$ are invariant under $(I-\lambda T)^{-1}$. It follows that

$$(I-\lambda T^{\mathfrak{R}})^{-1} = I+\lambda S_\lambda^{\mathfrak{R}} \quad (0 < |\lambda-\lambda_0| \leqslant \delta),$$

where $S_\lambda^{\mathfrak{R}}$ is the restriction to $\mathfrak{R}$ of S_λ. By Lemma 3.3.7 (applied to $T^{\mathfrak{R}}$), $S_\lambda^{\mathfrak{R}}$ is a trace class operator and its trace is bounded near λ_0. We suppose that

$$|\tau(S_\lambda^{\mathfrak{R}})| \leqslant M \quad (0 < |\lambda-\lambda_0| \leqslant \delta). \tag{19}$$

Let $\varphi_1, \ldots, \varphi_m$ be an orthonormal basis of $\mathfrak{N}$, with respect to which $T^{\mathfrak{N}}$ has a superdiagonal $m \times m$ matrix $A = [a_{ij}]$. Since $I-\lambda_0 T^{\mathfrak{N}}$ is nilpotent, so is its (superdiagonal) matrix $I_m-\lambda_0 A$. Hence the diagonal coefficients $1-\lambda_0 a_{jj}$ in this matrix, which are also its eigenvalues, are all zero. Thus $\lambda_0 a_{jj} = 1$ $(j = 1, \ldots, m)$, and

$$\begin{aligned}(I-\lambda_0 T)\varphi_j &= \varphi_j-\lambda_0 T^{\mathfrak{N}}\varphi_j \\ &= \varphi_j-\sum_{i=1}^{j} \lambda_0 \; {}_{ij}\varphi_i \\ &= -\sum_{i=1}^{j-1} \lambda_0 a_{ij}\varphi_i.\end{aligned}$$

It follows that, if $\mathfrak{N}_j$ is the subspace spanned by $\varphi_1, \ldots, \varphi_j$ $(j = 1, \ldots, m)$ and $\mathfrak{N}_0 = (0)$, then

$$(I-\lambda_0 T)(\mathfrak{N}_j) \subseteq \mathfrak{N}_{j-1} \quad (j = 1, \ldots, m).$$

Since $\mathfrak{R}$ is invariant under T,

$$(I-\lambda_0 T)(\mathfrak{N}_j+\mathfrak{R}) \subseteq \mathfrak{N}_{j-1}+\mathfrak{R} \quad (j = 1, \ldots, m). \tag{20}$$

The subspace $\mathfrak{N}_j+\mathfrak{R}$ is closed, by Theorem 1.1.1; it is spanned by φ_j and $\mathfrak{N}_{j-1}+\mathfrak{R}$, and $\varphi_j \notin \mathfrak{N}_{j-1}+\mathfrak{R}$ (otherwise,

$$\varphi_j = \alpha_1\varphi_1+\ldots+\alpha_{j-1}\varphi_{j-1}+x$$

for suitable scalars $\alpha_1, \ldots, \alpha_{j-1}$ and some x in $\mathfrak{R}$, and x is a non-zero vector in $\mathfrak{N}\cap\mathfrak{R}$, contradicting (18)). Hence we may choose a unit vector ψ_j in $\mathfrak{N}_j+\mathfrak{R}$ which is orthogonal to $\mathfrak{N}_{j-1}+\mathfrak{R}$; $\mathfrak{N}_j+\mathfrak{R}$ is spanned by ψ_j and $\mathfrak{N}_{j-1}+\mathfrak{R}$, so $\mathfrak{H}(=\mathfrak{N}+\mathfrak{R} = \mathfrak{N}_m+\mathfrak{R})$ is spanned by $\psi_1, \ldots, \psi_m$ and $\mathfrak{R}$. Thus, if $\{\theta_j : j \in J\}$ is an orthonormal basis of $\mathfrak{R}$, then $\{\psi_1, \ldots, \psi_m\} \cup \{\theta_j : j \in J\}$ is an orthonormal basis of $\mathfrak{H}$. It follows that

$$\begin{aligned}\tau(S_\lambda) &= \sum_{j=1}^{m} \langle S_\lambda\psi_j, \psi_j\rangle+\sum_{j\in J} \langle S_\lambda\theta_j, \theta_j\rangle \\ &= \sum_{j=1}^{m} \langle S_\lambda\psi_j, \psi_j\rangle+\sum_{j\in J} \langle S_\lambda^{\mathfrak{R}}\theta_j, \theta_j\rangle;\end{aligned}$$

that is,

$$\tau(S_\lambda) = \sum_{j=1}^{m} \langle S_\lambda\psi_j, \psi_j\rangle+\tau(S_\lambda^{\mathfrak{R}}) \quad (0 < |\lambda-\lambda_0| \leqslant \delta). \tag{21}$$

The next step is to compute $\langle S_\lambda\psi_j, \psi_j\rangle$. It follows from (20) that $\mathfrak{N}_j+\mathfrak{R}$ is invariant under T; as noted above, this implies that it is invariant under $(I-\lambda T)^{-1}$ for all complex numbers λ such that this inverse operator exists, and so invariant under S_λ when $0<|\lambda-\lambda_0|\leqslant\delta$. In particular, $S_\lambda\psi_j\in\mathfrak{N}_j+\mathfrak{R}$, so we may choose a complex number α and x in $\mathfrak{N}_{j-1}+\mathfrak{R}$ satisfying

$$S_\lambda\psi_j = \alpha\psi_j+x. \tag{22}$$

Then

$$\begin{aligned}\psi_j &= (I-\lambda T)(I+\lambda S_\lambda)\psi_j\\ &= (I-\lambda T)[(1+\lambda\alpha)\psi_j+\lambda x]\\ &= (1+\lambda\alpha)\left\{\left(1-\frac{\lambda}{\lambda_0}\right)I+\frac{\lambda}{\lambda_0}(I-\lambda_0 T)\right\}\psi_j+\lambda(I-\lambda T)x\\ &= (1+\lambda\alpha)\left(1-\frac{\lambda}{\lambda_0}\right)\psi_j+y,\end{aligned}$$

where

$$y = \frac{\lambda(1+\lambda\alpha)}{\lambda_0}(I-\lambda_0 T)\psi_j+\lambda(I-\lambda T)x.$$

Since $x\in\mathfrak{N}_{j-1}+\mathfrak{R}$ and this subspace is invariant under T, it follows from (20) that $y\in\mathfrak{N}_{j-1}+\mathfrak{R}$. However, $\psi_j\notin\mathfrak{N}_{j-1}+\mathfrak{R}$, so $y=0$,

$$(1+\lambda\alpha)\left(1-\frac{\lambda}{\lambda_0}\right) = 1,$$

and thus $\alpha=(\lambda_0-\lambda)^{-1}$. Since $x\in\mathfrak{N}_{j-1}+\mathfrak{R}$ and $\psi_j\in(\mathfrak{N}_{j-1}+\mathfrak{R})^\perp$, we can now deduce from (22) that

$$\langle S_\lambda\psi_j, \psi_j\rangle = \alpha = \frac{1}{\lambda_0-\lambda}.$$

This, with (21) yields

$$\tau(S_\lambda) = \frac{m}{\lambda_0-\lambda}+\tau(S_\lambda^{\mathfrak{R}}) \quad (0<|\lambda-\lambda_0|\leqslant\delta). \tag{23}$$

Comparison of (17) and (23), noting that $h(\lambda)$ and $\tau(S_\lambda^{\mathfrak{R}})$ are both bounded near λ_0, shows that $m=k$.

We now summarise, in the form of a theorem, the main results so far obtained in §3.3.

THEOREM 3.3.9. *Let T be a trace class operator acting on a Hilbert space $\mathcal{H}$, and suppose that the Fredholm determinant $d(\lambda, T)$ and the first Fredholm minor $D_{\lambda,T}$ are defined by means of equations (3), (4), (5), (8) and (9). Then the series for $d(\lambda, T)$ and $D_{\lambda,T}$ converge for all complex λ (the latter with respect to the norm $\|\cdot\|_1$ on the trace class $\mathcal{C}_1$). Furthermore*

(i) if λ is a complex number such that $d(\lambda, T)\neq 0$, then $I-\lambda T$ has an inverse given by

$$(I-\lambda T)^{-1} = I+\lambda[d(\lambda, T)]^{-1}D_{\lambda,T},$$

(ii) if λ is a complex number such that $d(\lambda, T)=0$, then λ^{-1} is an eigenvalue of T and its algebraic multiplicity as such is equal to the multiplicity of λ as a zero of $d(\lambda, T)$.

We now obtain estimates for the rates of growth of $d(\lambda, T)$ and $D_{\lambda,T}$ as $|\lambda|\to\infty$. In the following theorem Γ_p is the constant introduced in 3.2(20).

THEOREM 3.3.10. *(i) Suppose that T is a trace class operator acting on a Hilbert space $\mathcal{H}$, $d(\lambda, T)$ is the Fredholm determinant of T, $D_{\lambda,T}$ is the first Fredholm minor and r is a positive integer. Let (μ_j) be the infinite sequence consisting of the non-zero*

eigenvalues of $(T^*T)^{1/2}$, *counted according to their algebraic multiplicities and followed by zeros if the number of such eigenvalues is finite. Then*

$$|d(\lambda, T)| \leqslant \exp(|\lambda| \sum_{j=r+1}^{\infty} \mu_j) \prod_{j=1}^{r} (1+|\lambda|\mu_j); \tag{24}$$

$$||d(\lambda, T)I+\lambda D_{\lambda,T}|| \leqslant \exp(\tfrac{1}{2}+|\lambda| \sum_{j=r+1}^{\infty} \mu_j \prod_{j=1}^{r} (1+|\lambda|\mu_j); \tag{25}$$

$$|d(\lambda, T)| \leqslant \exp(|\lambda| \; ||T||_1); \tag{26}$$

$$||d(\lambda, T)I+\lambda D_{\lambda,T}|| \leqslant \exp(\tfrac{1}{2}+|\lambda| \; ||T||_1). \tag{27}$$

(ii) Suppose, further, that $0 < p < 1$ *and* $\Sigma \mu_j^p$ *converges. Then*

$$|d(\lambda, T)| \leqslant \exp(\Gamma_p|\lambda|^p \sum_{j=r+1}^{\infty} \mu_j^p) \prod_{j=1}^{r} (1+|\lambda|\mu_j); \tag{28}$$

$$||d(\lambda, T)I+\lambda D_{\lambda,T}|| \leqslant \exp(\tfrac{1}{2}+\Gamma_p|\lambda|^p \sum_{j=r+1}^{\infty} \mu_j^p) \prod_{j=1}^{r} (1+|\lambda|\mu_j); \tag{29}$$

$$|d(\lambda, T)| \leqslant \exp(\Gamma_p|\lambda|^p \sum_{j=1}^{\infty} \mu_j^p); \tag{30}$$

$$||d(\lambda, T)I+\lambda D_{\lambda,T}|| \leqslant \exp(\tfrac{1}{2}+\Gamma_p|\lambda|^p \sum_{j=1}^{\infty} \mu_j^p). \tag{31}$$

Proof. Suppose n is an integer such that $n \geqslant r$. By 3.2(23) and 3.2(28) we have

$$|d(\lambda, T_n)| \leqslant \exp(|\lambda| \sum_{j=r+1}^{n} \mu_j) \prod_{j=1}^{r} (1+|\lambda|\mu_j),$$

$$||d(\lambda, T_n)I+\lambda D_{\lambda,T_n}|| \leqslant \exp(\tfrac{1}{2}+|\lambda| \sum_{j=r+1}^{n} \mu_j) \prod_{j=1}^{r} (1+|\lambda|\mu_j).$$

By taking limits as $n \to \infty$ and appealing to Lemma 3.3.3(i) and (iii), we obtain (24) and (25).

The proofs of (26), ..., (31) are similar; in each case we quote the corresponding result for T_n, which has already been proved as part of Lemma 3.2.6 or Lemma 3.2.7 For (26) and (27) we also make use of (7).

COROLLARY 3.3.11. *Suppose that* $\epsilon > 0$ *and that the hypotheses of Theorem 3.3.10(i) are satisfied. Then there is a constant* M *such that, for all complex* λ,

$$|d(\lambda, T)| \leqslant M \exp(\epsilon|\lambda|), \tag{32}$$

$$||d(\lambda, T)I+\lambda D_{\lambda,T}|| \leqslant M \exp(\epsilon|\lambda|). \tag{33}$$

Suppose, further, that $0 < p < 1$ *and* $\Sigma \mu_j^p$ *converges. Then there is a constant* N *such that, for all complex* λ,

$$|d(\lambda, T)| \leqslant N \exp(\epsilon|\lambda|^p), \tag{34}$$

$$||d(\lambda, T)I+\lambda D_{\lambda,T}|| \leqslant N \exp(\epsilon|\lambda|^p). \tag{35}$$

Proof. Since $\Sigma \mu_j = ||T||_1 < \infty$, we can choose an integer r such that

$$\sum_{j=r+1}^{\infty} \mu_j < \tfrac{1}{2}\epsilon. \tag{36}$$

If M is defined to be the supremum, for $t \geqslant 0$, of the function

$$\exp(\tfrac{1}{2}-\tfrac{1}{2}\epsilon t) \prod_{j=1}^{r} (1+t\mu_j),$$

we have

$$\exp(\tfrac{1}{2}) \prod_{j=1}^{r} (1+t\mu_j) \leqslant M \exp(\tfrac{1}{2}\epsilon t) \quad (t \geqslant 0). \tag{37}$$

From (36) and (37) it follows that, for all complex λ,

$$\exp(\tfrac{1}{2}+|\lambda| \sum_{j=r+1}^{\infty} \mu_j) \prod_{j=1}^{r} (1+|\lambda|\mu_j) \leqslant M \exp(\epsilon|\lambda|).$$

With the aid of this last inequality, (32) and (33) are immediate consequences of (24) and (25), respectively.

If $\Sigma\, \mu_j^p < \infty$ we can choose an integer r such that

$$\Gamma_p \sum_{j=r+1}^{\infty} \mu_j^p < \tfrac{1}{2}\epsilon,$$

and define N to be the supremum, for $t \geqslant 0$, of the function

$$\exp(\tfrac{1}{2}-\tfrac{1}{2}\epsilon t^p) \prod_{j=1}^{r} (1+t\mu_j).$$

An argument similar to the proof of (32) and (33) then shows that (34) and (35) are immediate consequences of (28) and (29) respectively.

Note that the inequality (32) implies that $d(\lambda, T)$ is an entire function of finite order at most 1, and that if its order is exactly 1 then it is of minimum type (for the terminology used here, see §1.4). When $\Sigma\, \mu_j^p$ converges, (34) shows that $d(\lambda, T)$ has finite order at most p, and is of minimum type if its order is exactly p. Similar comments apply to the operator-valued entire function $D_{\lambda, T}$.

LEMMA 3.3.12. *Let T be a quasi-nilpotent trace class operator acting on a Hilbert space $\mathcal{H}$. Then $d(\lambda, T) = 1$ for all complex λ, and $\tau(T) = 0$.*

Proof. Since T is quasi-nilpotent, it follows from Theorem 3.3.9 that the entire function $d(\lambda, T)$ has no zeros; and, by (32), it has finite order at most 1. The Hadamard factorization theorem (Theorem 1.4.2) shows that $d(\lambda, T)$ has the form $\exp(\alpha+\beta\lambda)$, where α and β are constants. Since

$$\exp\, \alpha = d(0, T) = 1,$$

$d(\lambda, T) = \exp(\beta\lambda)$. Given any positive ϵ, it follows from (32) that $|\beta| < \epsilon$, so $\beta = 0$. Thus, for all complex λ,

$$1 = d(\lambda, T) = 1+\lambda d_1(T)+\ldots;$$

whence, by comparing coefficients and using (4), we have $\tau(T) = -d_1(T) = 0$.

It follows from Corollary 2.3.6, with $p = 1$, that the eigenvalues of a trace class operator T form an absolutely convergent series. We now show that the sum of this series is $\tau(T)$.

THEOREM 3.3.13. *Suppose that T is a trace class operator acting on a Hilbert space $\mathcal{H}$, and (λ_j) is the sequence of non-zero eigenvalues of T, counted according to their algebraic multiplicities. Then*

$$\tau(T) = \sum_j \lambda_j,$$

and the Fredholm determinant $d(\lambda, T)$ is given by

$$d(\lambda, T) = \prod_j (1-\lambda\lambda_j).$$

Proof. Let $\mathcal{M}$ be the closed subspace of $\mathcal{H}$ which is generated by the principal vectors of T^*, associated with non-zero eigenvalues, and let $\mathcal{N} = \mathcal{M}^\perp$. By Theorem 1.8.5, $\mathcal{N}$ is invariant under T and the restriction $T^{\mathcal{N}}$ of T to $\mathcal{N}$ is quasi-nilpotent. Since any orthonormal system in $\mathcal{N}$ is also an orthonormal system in $\mathcal{H}$, it follows at once from Definition 2.1.1 that $T^{\mathcal{N}}$ is a trace class operator on $\mathcal{N}$. By Lemma 3.3.12, $\tau(T^{\mathcal{N}}) = 0$.

The sequence of non-zero eigenvalues of T^*, with the correct multiplicities, is $(\bar{\lambda}_j)$. By Theorem 1.8.6 there is an orthonormal basis $\{\varphi_j\}$ of $\mathcal{M}$ such that $\langle T^*\varphi_j, \varphi_j\rangle = \bar{\lambda}_j$, whence $\langle T\varphi_j, \varphi_j\rangle = \lambda_j$.

Let $\{\psi_k\}$ be any orthonormal basis of $\mathfrak{N}$. Then $\{\varphi_j\} \cup \{\psi_k\}$ is an orthonormal basis of $\mathcal{H}$, and

$$\begin{aligned}\tau(T) &= \Sigma \langle T\varphi_j, \varphi_j\rangle + \Sigma \langle T\psi_k, \psi_k\rangle \\ &= \Sigma \lambda_j + \tau(T^{\mathfrak{N}}) \\ &= \Sigma \lambda_j.\end{aligned}$$

By Theorem 3.3.9, (λ_j^{-1}) is the sequence of zeros of the entire function $d(\lambda, T)$, with the correct multiplicities. Since $d(\lambda, T)$ has finite order not greater than 1, and $\Sigma |\lambda_j| < \infty$ (Corollary 2.3.6), the Hadamard factorization theorem (Theorem 1.4.2) shows that, for suitable constants α and β,

$$d(\lambda, T) = \exp(\alpha+\beta\lambda)\, P(\lambda), \tag{38}$$

where

$$P(\lambda) = \prod_j (1-\lambda\lambda_j).$$

Since $\Sigma |\lambda_j| < \infty$, the partial products

$$P_n(\lambda) = \prod_{j=1}^{n} (1-\lambda\lambda_j)$$

converge to $P(\lambda)$, uniformly in any bounded subset of the complex plane [12: §5.32, p.104]. It follows [12: §5.13, p. 95] that

$$\begin{aligned}P'(0) &= \lim_{n\to\infty} P_n'(0) \\ &= \lim_{n\to\infty} [-\sum_{j=1}^{n} \lambda_j] \\ &= -\sum_j \lambda_j \\ &= -\tau(T).\end{aligned}$$

When $\lambda = 0$, (38) gives $1 = \exp \alpha$, so

$$d(\lambda, T) = \exp(\beta\lambda) P(\lambda). \tag{39}$$

Thus

$$d'(\lambda, T) = \beta d(\lambda, T) + \exp(\beta\lambda) P'(\lambda),$$

and

$$\begin{aligned}\tau(T) &= -d_1(T) = -d'(0, T) \\ &= -\beta - P'(0) \\ &= -\beta + \tau(T).\end{aligned}$$

Hence $\beta = 0$ and, by (39),

$$d(\lambda, T) = P(\lambda) = \prod_j (1-\lambda\lambda_j).$$

3.4. *The resolvent of a quasi-nilpotent operator*

Suppose that T is a quasi-nilpotent operator acting on a Hilbert space $\mathcal{H}$. Then φ, defined by

$$\varphi(\lambda) = (I-\lambda T)^{-1},$$

is an operator-valued entire function. Our main purpose in this section, achieved in Theorem 3.4.6, is to estimate the rate of growth of $\|\varphi(\lambda)\|$ as $|\lambda| \to \infty$, given that T is in the von Neumann-Schatten class $\mathcal{C}_p$. When $\mathcal{H}$ is finite-dimensional the problem is a trivial one; accordingly, to avoid exceptional cases, we assume throughout §3.4 that $\mathcal{H}$ is infinite-dimensional.

Up until now we have considered the space $\mathcal{C}_p$ only for $p \geqslant 1$, but the work in this section requires the introduction of a class $\mathcal{C}_p$ for each positive p. Suppose that T is a compact linear operator acting on $\mathcal{H}$. Then T has a polar decomposition $T = V_T H_T$, where H_T $(= V_T^* T = (T^*T)^{1/2})$ is a positive compact operator. We shall

denote by $(\mu_j(T))$ the sequence of non-zero eigenvalues of H_T, arranged in decreasing order and counted according to their multiplicities. When T has finite rank n, the same is true of H_T, and so H_T has only n non-zero eigenvalues; in this case, it is convenient to define $\mu_j(T)$ to be 0 when $j > n$. With this convention, if T is any compact linear operator on $\mathcal{H}$, then $(\mu_j(T))$ is a decreasing real sequence with limit 0; and Theorem 1.9.3 shows that there exist orthonormal sequences (φ_j), (ψ_j) such that

$$Tx = \sum_j \mu_j \langle x, \varphi_j\rangle \psi_j \quad (x \in \mathcal{H}), \tag{1}$$

with $\mu_j = \mu_j(T)$.

DEFINITION 3.4.1. Suppose $p > 0$ and $\mathcal{H}$ is an infinite-dimensional Hilbert space. Then $\mathcal{C}_p$ is the set of all compact linear operators T on $\mathcal{H}$ such that

$$\sum_j \mu_j(T)^p < \infty.$$

For T in $\mathcal{C}_p$,

$$||T||_p = [\sum_j \mu_j(T)^p]^{1/p}.$$

From the final statement of Theorem 2.1.6, the remarks preceding Definition 2.3.2, and equation 2.3(4), it follows that the above definition is consistent with the terminology used, when $p \geqslant 1$, in Chapter 2.

Clearly, $\mathcal{C}_p \subseteq \mathcal{C}_q$ when $q > p > 0$. It is not difficult to show that $\mathcal{C}_p$ is a two-sided ideal in $\mathcal{B}(\mathcal{H})$, and contains the adjoint of each of its members; however, $||\ ||_p$ is not a norm on $\mathcal{C}_p$ when $0 < p < 1$. We omit the proofs, since we shall not make use of these results. The only additional property of $\mathcal{C}_p$ required below is a weakened form of Hölder's inequality. This is given in Lemma 3.4.4, after some auxiliary results.

LEMMA 3.4.2. *Suppose that T is a compact linear operator acting on a Hilbert space $\mathcal{H}$, and n is a positive integer. Then*

$$||T-F|| \geqslant \mu_n(T) \tag{2}$$

for each operator F on $\mathcal{H}$ with finite rank less than n. Equality occurs in (2) for at least one such F.

Proof. We may suppose that T is given by equation (1) and that, for suitable y_j and z_j in $\mathcal{H}(1 \leqslant j < n)$

$$Fx = \sum_{j=1}^{n-1} \langle x, y_j\rangle z_j \quad (x \in \mathcal{H}).$$

There exist complex numbers $\alpha_1, \ldots, \alpha_n$ such that

$$\sum_{i=1}^{n} |\alpha_i|^2 = 1, \quad \sum_{i=1}^{n} \alpha_i\langle \varphi_i, y_j\rangle = 0 \quad (1 \leqslant j < n).$$

If x is the unit vector $\sum \alpha_i\varphi_i$, then $\langle x, y_j\rangle = 0$ $(1 \leqslant j < n)$ and so $Fx = 0$. Thus

$$\begin{aligned} ||T-F||^2 &\geqslant ||Tx-Fx||^2 \\ &= ||Tx||^2 \\ &= ||\sum_{i=1}^{n} \mu_i(T)\alpha_i\psi_i||^2 \\ &= \sum_{i=1}^{n} \mu_i(T)^2|\alpha_i|^2 \\ &\geqslant \mu_n(T)^2 \sum_{i=1}^{n} |\alpha_i|^2 = \mu_n(T)^2, \end{aligned}$$

and $||T-F|| \geqslant \mu_n(T)$. It is easily verified that equality occurs when

$$Fx = \sum_{j=1}^{n-1} \mu_j \langle x, \varphi_j\rangle \psi_j \quad (x \in \mathcal{H}).$$

LEMMA 3.4.3. *Suppose that R and S are compact linear operators acting on a Hilbert space $\mathcal{H}$ and m, n are positive integers. Then*

$$\mu_{m+n-1}(RS) \leqslant \mu_m(R)\mu_n(S).$$

Proof. By Lemma 3.4.2 there exist operators R_1 (with finite rank not greater than $m-1$) and S_1 (with finite rank not greater than $n-1$) such that

$$||R-R_1|| = \mu_m(R), \quad ||S-S_1|| = \mu_n(S).$$

The operator F, defined by

$$F = R_1(S-S_1)+RS_1,$$

has finite rank not greater than $m+n-2$. It follows from Lemma 3.4.2 that

$$\begin{aligned}\mu_{m+n-1}(RS) &\leqslant ||RS-F|| \\ &= ||(R-R_1)(S-S_1)|| \\ &\leqslant ||R-R_1||\ ||S-S_1|| \\ &= \mu_m(R)\mu_n(S).\end{aligned}$$

The following result is of interest only when $t<1$, since Theorem 2.3.10 gives a sharper inequality when $t \geqslant 1$.

LEMMA 3.4.4. *Suppose that r, s, t are positive real numbers such that $r^{-1}+s^{-1} = t^{-1}$, and that $R \in \mathcal{C}_r$, $S \in \mathcal{C}_s$. Then $RS \in \mathcal{C}_t$ and $||RS||_t \leqslant 2^{1/t}||R||_r||S||_s$.*

Proof. Let $a = rt^{-1}$, $b = st^{-1}$, so that $a^{-1}+b^{-1} = 1$. It follows from Lemma 3.4.3 that

$$\mu_{2n}(RS) \leqslant \mu_{2n-1}(RS) \leqslant \mu_n(R)\mu_n(S) \quad (n = 1, 2, \ldots).$$

Hence

$$\begin{aligned}\sum_{j=1}^{\infty} \mu_j(RS)^t &= \sum_{n=1}^{\infty} \{\mu_{2n-1}(RS)^t+\mu_{2n}(RS)^t\} \\ &\leqslant 2\sum_{n=1}^{\infty} \mu_n(R)^t\mu_n(S)^t \\ &\leqslant 2[\sum_{n=1}^{\infty} \mu_n(R)^{ta}]^{1/a}[\sum_{n=1}^{\infty} \mu_n(S)^{tb}]^{1/b} \\ &= 2[\sum_{n=1}^{\infty} \mu_n(R)^r]^{t/r}[\sum_{n=1}^{\infty} \mu_n(S)^s]^{t/s}.\end{aligned}$$

Since the right-hand side is finite, $RS \in \mathcal{C}_t$; and, by taking tth. roots of both sides, we obtain the required inequality.

LEMMA 3.4.5. *Suppose that $T \in \mathcal{C}_p$, where $p \geqslant 1$, and let m be the smallest integer greater than p. Then*

$$T^m \in \mathcal{C}_{p/m}, \ ||T^m||_{p/m} \leqslant 2^{m/p}||T||_p^m.$$

Proof. We first prove, by induction on j, that

(3) $$T^j \in \mathcal{C}_{p/j}, \ ||T^j||_{p/j} \leqslant ||T||_p^j$$

for $j = 1, 2, \ldots, m-1$. When $j = 1$, there is nothing to prove. Suppose that $2 \leqslant k < m$, and that (3) is satisfied when $j = k-1$. By Theorem 2.3.10, with $r = p/(k-1)$, $s = p$ and $t = p/k$ ($\geqslant 1$),

$$T^k = T^{k-1}T \in \mathcal{C}_{p/k},$$

$$\begin{aligned}||T^k||_{p/k} &\leqslant ||T^{k-1}||_{p/(k-1)}||T||_p \\ &\leqslant ||T||_p^{k-1}||T||_p = ||T||_p^k.\end{aligned}$$

This completes the inductive proof of (3), when $1 \leqslant j < m$.

We now apply (3) (with $j = m-1$) and Lemma 3.4.4 (with $r = p/(m-1)$, $s = p$, $t = p/m$) to obtain

$$T^m = T^{m-1}T \in \mathcal{C}_{p/m},$$

$$||T^m||_{p/m} \leqslant 2^{m/p}||T^{m-1}||_{p/(m-1)}||T||_p$$
$$\leqslant 2^{m/p}||T||_p^m.$$

THEOREM 3.4.6. *Suppose that $p > 0$ and T is a quasi-nilpotent operator of class $\mathcal{C}_p$ acting on a Hilbert space $\mathcal{H}$. Then*

(i) there exist constants Γ_p and M_p (depending only on p) such that, for all complex λ,

(4) $$||(I-\lambda T)^{-1}|| \leqslant M_p \exp(\Gamma_p|\lambda|^p||T||_p^p);$$

(ii) given any positive ϵ, there is a constant M (which depends on p, T and ϵ) such that, for all complex λ,

(5) $$||(I-\lambda T)^{-1}|| \leqslant M \exp(\epsilon|\lambda|^p).$$

Proof. We consider separately two cases.

Case 1. $0 < p < 1$. Since $\mathcal{C}_p \subseteq \mathcal{C}_1$, T is a trace class operator, and the theory developed in §3.3 is available. By Lemma 3.3.12 and Theorem 3.3.9, $d(\lambda, T) = 1$ and

(6) $$(I-\lambda T)^{-1} = I+\lambda D_{\lambda,T} = d(\lambda, T)I+\lambda D_{\lambda,T},$$

for all complex λ. By (6) and 3.3(31),

$$||(I-\lambda T)^{-1}|| \leqslant \exp(\tfrac{1}{2}+\Gamma_p|\lambda|^p\Sigma\, \mu_j(T)^p)$$
$$= \exp(\tfrac{1}{2}) \exp(\Gamma_p|\lambda|^p||T||_p^p)$$

for all complex λ, where Γ_p is defined by 3.2(20) and depends only on p. By (6) and 3.3(35), there is a constant N such that

$$||(I-\lambda T)^{-1}|| \leqslant N \exp(\epsilon|\lambda|^p)$$

for all complex λ. This completes the proof of the theorem in case 1.

Case 2. $p \geqslant 1$. Let m be the smallest integer greater than p. By Lemma 3.4.5,

(7) $$T^m \in \mathcal{C}_q, \qquad ||T^m||_q \leqslant 2^{1/q}||T||_p^m,$$

where $q = p/m$. Since $0 < q < 1$ and T^m is a quasi-nilpotent element of $\mathcal{C}_q$, the results established in case 1 are applicable to T^m. It follows that, for all complex λ,

$$||(I-\lambda^m T^m)^{-1}|| \leqslant \exp(\tfrac{1}{2}+\Gamma_q|\lambda^m|^q||T^m||_q^q)$$
$$\leqslant \exp(\tfrac{1}{2}+2\Gamma_q|\lambda|^{mq}||T||_p^{mq}),$$

(8) $$||(I-\lambda^m T^m)^{-1}|| \leqslant \exp(\tfrac{1}{2}+2\Gamma_q|\lambda|^p||T||_p^p).$$

Furthermore, there is a constant N such that, for all complex λ,

(9) $$||(I-\lambda^m T^m)^{-1}|| \leqslant N \exp(\tfrac{1}{2}\epsilon|\lambda^m|^q) = N \exp(\tfrac{1}{2}\epsilon|\lambda|^p).$$

Since

$$(I-\lambda^m T^m) = (I+\lambda T+\lambda^2T^2+\ldots+\lambda^{m-1}T^{m-1})(I-\lambda T),$$

we have

$$(I-\lambda T)^{-1} = (I+\lambda T+\lambda^2T^2+\ldots+\lambda^{m-1}T^{m-1})(I-\lambda^m T^m)^{-1}.$$

From this last identity and the fact that $||T|| \leqslant ||T||_p$, it follows that

(10) $$||(I-\lambda T)^{-1}|| \leqslant (1+|\lambda|\ ||T||_p+|\lambda|^2||T||_p^2+\ldots+ +|\lambda|^{m-1}||T||_p^{m-1})||(I-\lambda^m T^m)^{-1}||.$$

There exist constants K_p (depending only on p) and K such that, for all positive t,

(11) $$1+t+t^2+\ldots+t^{m-1} \leqslant K_p\exp(t^p),$$

(12) $$1+t||T||_p+t^2||T||_p^2+\ldots+t^{m-1}||T||_p^{m-1} \leqslant K \exp(\tfrac{1}{2}\epsilon t^p).$$

From (10), (11) and (8),

$$||(I-\lambda T)^{-1}|| \leqslant K_p \exp(|\lambda|^p ||T||_p^p) \exp(\tfrac{1}{2}+2\Gamma_q |\lambda|^p ||T||_p^p)$$
$$= M_p \exp(\Gamma_p |\lambda|^p ||T||_p^p),$$

where $M_p = K_p \exp(\tfrac{1}{2})$ and $\Gamma_p = 2\Gamma_q+1$. By (10, (12) and (9),

$$||(I-\lambda T)^{-1}|| \leqslant K \exp(\tfrac{1}{2}\epsilon|\lambda|^p) N \exp(\tfrac{1}{2}\epsilon|\lambda|^p)$$
$$= M \exp(\epsilon|\lambda|^p),$$

where $M = KN$. This completes the proof of Theorem 3.4.6.

Theorem 3.4.6(ii) shows that, if $\varphi(\lambda) = (I-\lambda T)^{-1}$ for some quasi-nilpotent operator T of class $\mathcal{C}_p$, where $0 < p < \infty$, then the operator-valued entire function φ has finite order at most p, and is of minimum type if its order is exactly p.

3.5. *Some applications*

Suppose that T is a compact linear operator acting on a Hilbert space $\mathcal{H}$, and let $\mathfrak{N}_T$ be the null space of T and $\mathfrak{R}_T$ the closed range space of T; that is, $\mathfrak{N}_T = \{x \in \mathcal{H} : Tx = 0\}$, while $\mathfrak{R}_T$ is the closure of $\{Tx : x \in \mathcal{H}\}$. If x is a principal vector of T, associated with a non-zero eigenvalue μ, then $(T-\mu I)^n x = 0$ for some positive integer n, and from this it follows that $x \in \mathfrak{R}_T$. Hence

$$\mathcal{P}_T \subseteq \mathfrak{R}_T, \tag{1}$$

where $\mathcal{P}_T$ denotes the closed subspace generated by the set of all principal vectors of T associated with non-zero eigenvalues. In this section, we are interested in finding conditions under which $\mathcal{P}_T = \mathfrak{R}_T$; that is, conditions under which, for each x in $\mathcal{H}$, Tx can be approximated by linear combinations of principal vectors.

Since $\mathfrak{R}_T^\perp = \mathfrak{N}_{T^*}$, it follows from (1) that

$$\mathcal{P}_T^\perp \supseteq \mathfrak{N}_{T^*}, \tag{2}$$

with equality if and only if $\mathcal{P}_T = \mathfrak{R}_T$.

From the foregoing discussion, we see that $\mathcal{P}_T = \mathfrak{R}_T$ if and only if $T^*x = 0$ for each x in $\mathcal{P}_T^\perp$. If T satisfies the stronger condition, $Tx = T^*x = 0$ for each x in $\mathcal{P}_T^\perp$, then $\mathcal{P}_T = \mathfrak{R}_T$ and, in addition, the closed subspace generated by the principal vectors (associated with non-zero eigenvalues) and null space of T is the whole of $\mathcal{H}$.

We recall from Theorem 1.8.5 that $\mathcal{P}_T^\perp$ is invariant under T^*, and the restriction N of T^* to $\mathcal{P}_T^\perp$ is quasi-nilpotent. If $1 \leqslant p < \infty$ and $T \in \mathcal{C}_p$, then the self-adjoint part $A = \tfrac{1}{2}(T+T^*)$, the skew-adjoint part $B = \tfrac{1}{2i}(T^*-T)$ and the adjoint T^* of T all lie in $\mathcal{C}_p$; since any orthonormal system in $\mathcal{P}_T^\perp$ is also an orthonormal system in $\mathcal{H}$, it follows at once from Definition 2.1.1 that $N \in \mathcal{C}_p$.

The following result is due to V. B. Lidskiǐ [39].

THEOREM 3.5.1. *Suppose that T is a trace class operator acting on a Hilbert space $\mathcal{H}$, and $\tfrac{1}{2i}(T^*-T) \geqslant 0$. Then $\mathcal{P}_T = \mathfrak{R}_T$, and the closed subspace generated by the principal vectors (associated with non-zero eigenvalues) and null space of T is the whole of $\mathcal{H}$.*

Proof. By the preceding discussion, it is sufficient to show that $Tx = T^*x = 0$ whenever $x \in \mathcal{P}_T^\perp$. If $A = \tfrac{1}{2}(T+T^*)$ and $B = \tfrac{1}{2i}(T^*-T)$, then A and B are self-adjoint trace class operators, $B \geqslant 0$, $T = A+iB$ and $T^* = A-iB$. The restriction N of T^* to $\mathcal{P}_T^\perp$ is a quasi-nilpotent trace class operator and, by Lemma 3.3.12, $\tau(N) = 0$. If (ψ_j) is an orthonormal basis of $\mathcal{P}_T^\perp$, then

$$0 = \tau(N) = \Sigma \langle N\psi_j, \psi_j \rangle$$
$$= \Sigma \langle T^*\psi_j, \psi_j \rangle$$
$$= \Sigma \{\langle A\psi_j, \psi_j \rangle - i\langle B\psi_j, \psi_j \rangle\}.$$

Since $B \geqslant 0$, it follows by taking imaginary parts that $\langle B\psi_j, \psi_j\rangle = 0$ for each j. Thus $B\psi_j = 0$ and, since linear combination of ψ_j's are dense in $\mathcal{P}_T^{\perp}$,

(3) $$Bx = 0 \quad (x \in \mathcal{P}_T^{\perp}).$$

Hence

(4) $$Ax = (A-iB)x = T^*x = Nx \quad (x \in \mathcal{P}_T^{\perp}),$$

and $\langle Nx, y\rangle = \langle Ax, y\rangle = \langle x, Ay\rangle = \langle x, Ny\rangle$ whenever $x, y \in \mathcal{P}_T^{\perp}$. It follows that the quasi-nilpotent operator N is also self-adjoint; Corollary 1.7.5 shows that $N = 0$ and, by (4),

(5) $$Ax = 0 \quad (x \in \mathcal{P}_T^{\perp}).$$

By (3) and (5), $Tx = T^*x = 0$ for each x in $\mathcal{P}_T^{\perp}$.

The following theorem is due to V. B. Lidskiĭ [37, 38] and L. A. Sahnovic [56].

THEOREM 3.5.2. *Suppose that* $T \in \mathcal{C}_2$, $\frac{1}{2}(T+T^*) \geqslant 0$ *and* $\frac{1}{2}i(T^*-T) \geqslant 0$. *Then* $\mathcal{P}_T = \mathcal{R}_T$, *and the closed subspace generated by the principal vectors (associated with non-zero eigenvalues) and null space of* T *is the whole of the Hilbert space* $\mathcal{H}$ *on which* T *acts.*

Proof. Once again, it is sufficient to prove that $Tx = T^*x = 0$ whenever $x \in \mathcal{P}_T^{\perp}$. The restriction N of T^* to $\mathcal{P}_T^{\perp}$ is a quasi-nilpotent operator of class $\mathcal{C}_2$; so N^2 is a quasi-nilpotent trace class operator and, by Lemma 3.3.12,

(6) $$\tau(N^2) = 0.$$

Let $N = H-iK$, where H and K are self-adjoint operators on $\mathcal{P}_T^{\perp}$. Then H and K are of class $\mathcal{C}_2$, since $N \in \mathcal{C}_2$. For each x in $\mathcal{P}_T^{\perp}$,

$$\begin{aligned}\langle Hx, x\rangle - i\langle Kx, x\rangle &= \langle Nx, x\rangle \\ &= \langle T^*x, x\rangle \\ &= \langle Ax, x\rangle - i\langle Bx, x\rangle,\end{aligned}$$

where $A = \frac{1}{2}(T+T^*)$, $B = \frac{1}{2}i(T^*-T)$. Since A and B are positive operators, so are H and K. By (6),

$$0 = \tau(N^2) = \tau(H^2-K^2) - i\tau(HK+KH);$$

and, since $\tau(HK) = \tau(KH) = \tau((HK)^*) = \overline{\tau(HK)}$ by Theorem 2.3.13, it follows that

(7) $$\tau(HK) = 0.$$

We may suppose that

(8) $$Kx = \sum_j \mu_j \langle x, \varphi_j\rangle \varphi_j \quad (x \in \mathcal{P}_T^{\perp}),$$

where (φ_j) is an orthonormal sequence in $\mathcal{P}_T^{\perp}$ and (μ_j) is a decreasing sequence of positive real numbers. Let (ψ_k) be another orthonormal system, such that $(\varphi_j) \cup (\psi_k)$ is an orthonormal basis of $\mathcal{P}_T^{\perp}$. By (8), $K\psi_k = 0$ for each k; so

$$\begin{aligned}0 = \tau(HK) &= \sum_j \langle HK\varphi_j, \varphi_j\rangle + \sum_k \langle HK\psi_k, \psi_k\rangle \\ &= \sum_j \mu_j \langle H\varphi_j, \varphi_j\rangle.\end{aligned}$$

Since $\mu_j > 0$ and $H \geqslant 0$, it follows that $H\varphi_j = 0$ for each j. By (8), $HK = 0$; thus $KH = (HK)^* = 0 = HK$, and N $(= H-iK)$ is both normal and (as noted above) quasi-nilpotent. By Corollary 1.7.5, $N = 0$.

For each x in $\mathcal{P}_T^{\perp}$,

$$0 = \langle Nx, x\rangle = \langle Ax, x\rangle - i\langle Bx, x\rangle.$$

Since A and B are positive operators, it follows that $Ax = Bx = 0$; so $Tx = T^*x = 0$, and the theorem is proved.

CHAPTER 4

Superdiagonal Representation of Compact Linear Operators

4.1. *Introduction*

An $n \times n$ matrix $A = [a_{jk}]$ is said to be *superdiagonal* if $a_{jk} = 0$ whenever $j > k$; the eigenvalues of such a matrix are its diagonal elements $a_{11}, a_{22}, \ldots, a_{nn}$. An elementary theorem (see, for example, [26: p. 144]) asserts that, if B is an $n \times n$ complex matrix, then there is a unitary matrix U such that $U^{-1}BU$ is superdiagonal. This result can be reformulated in the following way [26: pp. 107, 144]; if T is a linear operator acting on an n-dimensional Hilbert space $\mathcal{H}$, then there exist subspaces $L_0, L_1, \ldots, L_n$ of $\mathcal{H}$ such that

(i) $\{0\} = L_0 \subseteq L_1 \subseteq \ldots \subseteq L_n = \mathcal{H}$,

(ii) L_j is j-dimensional ($j = 0, 1, \ldots, n$),

(iii) each L_j is invariant under T.

If, for each $j = 1, \ldots, n$, e_j is a unit vector in $L_j \cap L_{j-1}^{\perp}$, then $\{e_1, \ldots, e_n\}$ is an orthonormal basis of $\mathcal{H}$, with respect to which T has a superdiagonal matrix A. The eigenvalues of the operator T, which coincide with those of the matrix A, are therefore precisely the diagonal elements of A. It follows easily that the eigenvalues of A, counted according to their algebraic multiplicities, are the scalars $\lambda_1, \ldots, \lambda_n$ determined by the condition $Te_j - \lambda_j e_j \in L_{j-1}$ ($j = 1, \ldots, n$). Since $e_j \in L_{j-1}^{\perp}$, we deduce that

$$0 = \langle Te_j - \lambda_j e_j,\ e_j \rangle = \langle Te_j,\ e_j \rangle - \lambda_j,$$

and so

$$\lambda_j = \langle Te_j, e_j \rangle.$$

The operator T is nilpotent if and only if $\lambda_j = 0$ $(j = 1, \ldots, n)$, and this is equivalent to the assertion that $Te_j \in L_{j-1}$ $(j = 1, \ldots, n)$. Since L_j is the linear span of e_j and L_{j-1}, and $T(L_{j-1}) \subseteq L_{j-1}$, it follows that T is nilpotent if and only if $T(L_j) \subseteq L_{j-1}$ $(j = 1, \ldots, n)$.

The preceding discussion suggests that a family $\{L_j : j = 1, \ldots, n\}$ of subspaces, which has properties (i), (ii) and (iii) above, determines a 'superdiagonal representation' of T. In the present chapter, we consider the application of this idea to compact linear operators acting on infinite-dimensional spaces. In §4.2, we prove the theorem, due to Aronszajn and Smith [1], that a compact linear operator T acting on a complex Banach space X has a proper closed invariant subspace (this result has recently been extended in a number of ways; see [2, 3, 4, 13, 17, 27, 28]). In §4.3 the Aronszajn Smith theorem is used, with a simple Zorn's Lemma argument, to prove the existence of a maximal totally ordered family $\mathcal{F}$ of closed subspaces of X, each of which is invariant under T. It is then possible to obtain satisfactory analogues of the finite-dimensional results, described above, concerning the eigenvalues of T and the conditions under which T is quasi-nilpotent.

In §4.4 we consider the superdiagonal representation of compact linear operators acting on Hilbert spaces, together with the following question: if $\mathcal{F}$ is a maximal totally ordered family of closed subspaces of a Hilbert space $\mathcal{H}$ and S is a compact self-adjoint operator on $\mathcal{H}$, is there a quasi-nilpotent operator T on $\mathcal{H}$ which has skew-adjoint part S and leaves invariant each subspace in $\mathcal{F}$? In order to clarify the nature of this question, we now consider the problem in finite-dimensional spaces. In this case, we can suppose that the family $\mathcal{F}$ consists of subspaces $L_0, L_1, \ldots, L_n$ satisfying conditions (i) and (ii) above, and we can choose an orthonormal basis $e_1, \ldots, e_n$ of $\mathcal{H}$ such that $e_j \in L_j \cap L_{j-1}^{\perp}$ $(j = 1, \ldots, n)$. The given operator S has a hermitian matrix $[s_{jk}]$ with respect to this basis. If T is a nilpotent operator acting on $\mathcal{H}$, and leaves invariant each of the subspaces L_j, then its matrix $[t_{jk}]$ with respect to the given orthonormal basis is both superdiagonal and nilpotent, and therefore satisfies

$$(1) \qquad t_{jk} = 0 \quad (1 \leqslant k \leqslant j \leqslant n).$$

The required condition, that $S = \tfrac{1}{2}i(T^* - T)$, is equivalent to

$$(2) \qquad s_{jk} = \begin{cases} -\tfrac{1}{2}it_{jk} & (j < k), \\ 0 & (j = k), \\ \tfrac{1}{2}i\bar{t}_{k,j} & (j > k). \end{cases}$$

Since $[s_{jk}]$ is a hermitian matrix, a solution $[t_{jk}]$ of (1) and (2) exists if and only if

$$(3) \qquad s_{jj} = 0 \quad (j = 1, \ldots, n);$$

the solution is then unique, and is given by

$$(4) \qquad t_{jk} = \begin{cases} 2is_{jk} & (j < k), \\ 0 & (j \geqslant k). \end{cases}$$

It is not difficult to show that, if E_j denotes the projection from $\mathcal{H}$ onto L_j, then (3) and (4) can be rewritten in the form

$$(3') \qquad (E_j - E_{j-1})S(E_j - E_{j-1}) = 0 \quad (j = 1, \ldots, n),$$

$$(4') \qquad T = 2i \sum_{k=1}^{n} E_{k-1}S(E_k - E_{k-1}).$$

Our conclusion can therefore be stated as follows: the condition (3′) is both necessary and sufficient for the existence of a nilpotent operator T on $\mathcal{H}$ which has skew-adjoint part S and leaves each of the subspaces L_j invariant; when (3′) is satisfied, the operator T is unique, and is given by (4′).

In §4.4 we consider the problem, in the form set out at the beginning of the preceding paragraph, in the infinite-dimensional case. It turns out that there is a simple condition on S, analogous to (3′), which is *necessary* for the existence of an operator T with the requisite properties. Unfortunately, this condition is *not sufficient*; it can be strengthened, so as to become a sufficient condition, in various ways (for example, by requiring S to lie in a von Neumann–Schatten class $\mathcal{C}_p$, $p < \infty$). No simple necessary and sufficient condition for the existence of a suitable operator T is known. Given a general compact self-adjoint operator S, it is not difficult to show that, if such an operator T exists, then it is unique and is compact; furthermore, T can be expressed, in terms of S and $\mathcal{F}$, as a 'superdiagonal integral' analogous to the sum occurring in (4′). In view of equation (4) it is reasonable to regard T, when it exists, as the 'superdiagonal part' of S relative to the given totally ordered family $\mathcal{F}$ of subspaces of $\mathcal{H}$.

In §§4.5 and 4.6 we consider the integral operator V, acting on the space $L_2(0,1)$ (Lebesgue measure) which is defined by

$$(Vx)(s) = i\int_0^s x(t)dt \quad (x \in L_2(0,1) : 0 \leqslant s \leqslant 1). \tag{5}$$

If $0 \leqslant \lambda \leqslant 1$, then the closed subspace

$$L_\lambda = \{x \in L_2(0,1) : x(s) = 0 \text{ almost everywhere on } [0, \lambda]\} \tag{6}$$

of $L_2(0,1)$ is invariant under V. It turns out that there are three simple properties of V which characterize this operator up to unitary equivalence (that is, any operator with these properties is unitarily equivalent to V). Furthermore, the subspaces L_λ $(0 \leqslant \lambda \leqslant 1)$ are the only closed invariant subspaces of V. Thus V is an example of a *unicellular* operator; that is, an operator whose closed invariant subspaces are totally ordered by the inclusion relation $\subseteq$. In §4.5 we prove these results by using the theory of superdiagonal integrals developed in §4.4; an alternative, more elementary, treatment is given in §4.6.

The results described in §4.4 were originally obtained by various Russian mathematicians, notably Brodskii [6], Gohberg and Krein [18, 19, 20] and Macaev [44, 45]. The characterization (up to unitary equivalence) of the operator V is a special case of more general results due to Livsic [42]. The unicellularity of V was first proved by Dixmier [14]; subsequently, a number of proofs have been given [5, 15, 30, 8]. The treatment given in §4.5 is based on [8], and I am indebted to A. L. Brown for pointing out that the method used in [8] to establish unicellularity can easily be adapted to obtain the characterization of V up to unitary equivalence. The methods used in §4.6 are derived from [5] and [30, 31]. The unicellularity of certain other Volterra integral operators has been established by Kalisch [29].

4.2. *Invariant subspaces of compact linear operators*

Our purpose in this section is to prove the following result.

THEOREM 4.2.1. *If T is a compact linear operator acting on a complex Banach space X of dimension greater than 1, then T has a proper closed invariant subspace.*

The original proof of this result [1] has subsequently been simplified a little; the version given below is derived from [16].

Before embarking on the proof, we require some further notation and a number of auxiliary result.

Throughout §4.2, T is a compact linear operator acting on a complex Banach space X. In the case of finite-dimensional Banach spaces, the result of Theorem 4.2.1 is contained in the elementary linear algebra discussed at the beginning of §4.1; we may therefore assume that X is infinite-dimensional. The closed subspace of X generated by a (finite or countable) subset $\{x_1, x_2, \ldots\}$ will be denoted by $\mathfrak{L}(x_1, x_2, \ldots)$. Once and for all, we choose a unit vector e in X. The closed subspace $X_0 = \mathfrak{L}(e, Te, T^2e, \ldots)$ is invariant under T, and is not $\{0\}$. If $X_0 \neq X$, then X_0 is a proper closed invariant subspace of T; we may therefore assume that $X_0 = X$. Let

(1) $$X_n = \mathfrak{L}(e, Te, T^2e, \ldots, T^{n-1}e) \quad (n = 1, 2, \ldots).$$

If the vectors $e, Te, T^2e, \ldots$ are linearly dependent, there is a positive integer m such that $T^m e \in X_m$; from this, it follows easily that $T^n e \in X_m$ $(n = 0, 1, 2, \ldots)$, and X_m coincides with X_0 $(= X)$, contrary to our supposition that X is infinite-dimensional. Hence the vectors $e, Te, T^2e, \ldots$ are linearly independent, and the subspace X_n is n-dimensional. In view of the preceding discussion, we shall assume in future that

(2) the vectors e, Te, $T^2e, \ldots$ are linearly independent,

(3) $$X = \mathfrak{L}(e, Te, T^2e, \ldots).$$

If Z is a closed subspace of X, and $x \in X$, we write $d(x, Z)$ for the distance $\inf\{||x-z|| : z \in Z\}$ of x from Z. It is apparent that $d(x, Z) \leqslant ||x||$ and that, for a given subspace Z, the mapping $x \to d(x, Z)$ is a seminorm on X. It follows that, if (Z_k) is a sequence of closed subspaces of X, then the mapping

$$x \to \limsup_{k \to \infty} d(x, Z_k)$$

is a seminorm on X, majorized by $|| \;||$. This implies that the subset $\liminf Z_k$, defined by

(4) $$\liminf Z_k = \{x \in X : \limsup_{k \to \infty} d(x, Z_k) = 0\},$$

is a closed subspace of X.

If $x \in X$ and $\epsilon > 0$, it follows from (3) that there is a positive integer m and a linear combination z of $e, Te, T^2e, \ldots, T^{m-1}e$ such that $||x-z|| < \epsilon$; thus $z \in X_n$ and

$$d(x, X_n) \leqslant ||x-z|| < \epsilon,$$

whenever $n \geqslant m$. Hence

(5) $$\limsup_{n \to \infty} d(x, X_n) = 0 \quad \text{for each } x \text{ in } X;$$

equivalently, $\liminf X_n = X$.

For each positive integer n, let x_n be a vector satisfying

(6) $$x_n \in X_n, \quad ||T^n e - x_n|| < 2d(T^n e, X_n).$$

There is a unique linear operator T_n, acting on the space X_n, such that

$$T_n(T^k e) = T^{k+1}e \quad (k = 0, 1, \ldots, n-2),$$

$$T_n(T^{n-1}e) = x_n.$$

For each element $x = \alpha_0 e + \alpha_1 Te + \ldots + \alpha_{n-1}T^{n-1}e$ of X_n,

$$\begin{aligned} ||Tx - T_n x|| &= ||\sum_{k=0}^{n-1} \alpha_k[T^{k+1}e - T_n(T^k e)]|| \\ &= ||\alpha_{n-1}(T^n e - x_n)|| \\ &= |\alpha_{n-1}| \; ||T^n e - x_n|| \end{aligned}$$

$$\leqslant 2|\alpha_{n-1}|d(T^n e, X_n)$$
$$= 2d(\alpha_{n-1}T^n e, X_n).$$

Since

$$Tx-\alpha_{n-1}T^n e = \sum_{k=0}^{n-1} \alpha_k T^{k+1}e-\alpha_{n-1}T^n e$$
$$= \sum_{k=0}^{n-2} \alpha_k T^{k+1}e \in X_n,$$

it follows that $d(\alpha_{n-1}T^n e, X_n) = d(Tx, X_n)$; thus

(7) $$||Tx-T_n x|| \leqslant 2d(Tx, X_n) \quad (x \in X_n).$$

Since T_n is a linear operator acting on the n-dimensional vector space X_n, it follows from the elementary linear algebra discussed at the beginning of §4.1 that there exist subspaces $L_{m,n}$ $(0 \leqslant m \leqslant n)$ of X_n such that

(8) $$\{0\} = L_{0,n} \subseteq L_{1,n} \subseteq \ldots \subseteq L_{n,n} = X_n,$$

(9) $L_{m,n}$ is m-dimensional,

(10) $L_{m,n}$ is invariant under T_n.

The following lemma shows how invariant subspaces of T may be constructed from suitable sequences of the subspaces $L_{m,n}$.

LEMMA 4.2.2. *Suppose that* $(m(k))_{1\leqslant k<\infty}$ *and* $(n(k))_{1\leqslant k<\infty}$ *are sequences of integers such that* $0 \leqslant m(k) \leqslant n(k) < n(k+1)$ $(k = 1, 2, \ldots)$, *and* L *is the closed subspace* $\liminf L_{m(k),n(k)}$ *of* X.

(i) *If* $z \in X$, $x_k \in L_{m(k),n(k)}$ $(k = 1, 2, \ldots)$ *and* $||z-Tx_k|| \to 0$ *as* $k \to \infty$, *then* $z \in L$.

(ii) L *is invariant under* T.

Proof. Suppose that z and x_k satisfy the conditions set out in (i). Since $L_{m(k),n(k)}$ is invariant under $T_{n(k)}$, it follows that $T_{n(k)}x_k \in L_{m(k),n(k)}$; thus

$$d(z, L_{m(k),n(k)}) \leqslant ||z-T_{n(k)}x_k||$$
$$\leqslant ||z-Tx_k||+||Tx_k-T_{n(k)}x_k||.$$

This, together with (7), gives

$$d(z, L_{m(k),n(k)}) \leqslant ||z-Tx_k||+2d(Tx_k, X_{n(k)})$$
$$\leqslant ||z-Tx_k||+2d(Tx_k-z, X_{n(k)})+2d(z, X_{n(k)})$$
$$\leqslant 3||z-Tx_k||+2d(z, X_{n(k)})$$
$$\to 0$$

as $k \to \infty$, by (5). Hence $z \in \liminf L_{m(k),n(k)} = L$, and (i) is proved

If $x \in L$ then $d(x, L_{m(k),n(k)}) \to 0$ as $k \to \infty$. We may choose x_k in $L_{m(k),n(k)}$ such that $||x-x_k|| \leqslant 2d(x, L_{m(k),n(k)})$, and then

$$||Tx-Tx_k|| \leqslant ||T||\ ||x-x_k|| \leqslant 2||T||d(x, L_{m(k),n(k)}) \to 0.$$

Part (i) of the lemma, with Tx in place of z, implies that $Tx \in L$. This proves (ii).

The remainder of this section is devoted to showing that the sequences $(m(k))$ and $(n(k))$ can be chosen so that the subspace L occuring in Lemma 4.2.2 is neither $\{0\}$ nor X. Suppose that α is a real number satisfying

(11) $$0 < \alpha < ||Te||/||T||.$$

Clearly $\alpha < ||e|| = 1$; furthermore, $e \in X_n = L_{n,n}$ and, by (8),

$$1 = ||e|| = d(e,L_{0,n}) \geqslant d(e, L_{1,n}) \geqslant \ldots \geqslant d(e,L_{n,n}) = 0.$$

It follows that, for each $n = 1, 2, 3, \ldots$, there is an integer $m(n)$ such that

$$(12) \qquad 0 \leqslant m(n) < n,\ d(e, L_{m(n),n}) \geqslant \alpha > d(e, L_{m(n)+1,n}).$$

LEMMA 4.2.3. *If $(n(k))$ is a strictly increasing sequence of positive integers, then* $\liminf L_{m(n(k)),n(k)} \neq X$.

Proof. Since $d(e, L_{m(n(k)),n(k)}) \geqslant \alpha > 0$ $(k = 1, 2, \ldots)$, by (12), it follows that $e \notin \liminf L_{m(n(k)),n(k)}$.

If there is a strictly increasing sequence $(\ell(k))$ of positive integers such that $\liminf L_{m(\ell(k)),\ell(k)} \neq \{0\}$, it follows from Lemmas 4.2.3 and 4.2.2 that this subspace is not the whole of X, and is therefore a proper closed invariant subspace of T. It remains to deal with the case in which no such sequence $(\ell(k))$ exists.

LEMMA 4.2.4. *Suppose that* $\liminf L_{m(\ell(k)),\ell(k)} = \{0\}$ *whenever $(\ell(k))$ is a strictly increasing sequence of positive integers. Suppose also that $(n(k))$ is a strictly increasing sequence of positive integers, and (x_k) is a bounded sequence of vectors such that $x_k \in L_{m(n(k)),n(k)}$ $(k = 1, 2, 3, \ldots)$. Then $||Tx_k|| \to 0$ as $k \to \infty$.*

Proof. Suppose the contrary. By passing to a subsequence of $(n(k))$, we may suppose that there is a positive real number δ such that $||Tx_k|| \geqslant \delta$ $(k = 1, 2, \ldots)$. Since T is compact, we may assume also that (Tx_k) converges to an element z of X. Then $||z|| \geqslant \delta$ and, by Lemma 4.2.2, $z \in \liminf L_{m(n(k)),n(k)}$; so $\liminf L_{m(n(k)),n(k)} \neq \{0\}$, contradicting the hypothesis of the lemma.

Proof of Theorem 4.2.1. Since the logical structure of the proof is rather complicated, we begin with a résumé of our progress so far. We have seen that it is sufficient to consider the case in which (2) and (3) are satisfied. In this case, we have defined subspaces X_n $(n = 1, 2, \ldots)$ by (1), and have introduced the operators T_n and subspaces $L_{m,n}$ $(0 \leqslant m \leqslant n;\ n = 1, 2, \ldots)$ satisfying (8), (9) and (10). In the discussion preceding Lemma 4.2.4 we have shown that, if there is any strictly increasing sequence $(\ell(k))$ of positive integers such that $\liminf L_{m(\ell(k)),\ell(k)} \neq \{0\}$, then this last subspace is a proper closed invariant subspace of T. Henceforth we consider the case in which no such sequence $(\ell(k))$ exists; in this case, the hypothesis of Lemma 4.2.4 is satisfied, and so the conclusion of that lemma is available.

It follows from (12) that, for each positive integer n, there is a vector x_n such that

$$(13) \qquad x_n \in L_{m(n)+1,n}, \qquad ||e - x_n|| < \alpha.$$

Since T is compact and the sequence (x_n) is bounded, there is a strictly increasing sequence $(n(k))$ of positive integers such that $(Tx_{n(k)})$ converges to an element z of X. By Lemma 4.2.2(ii), the closed subspace $L = \liminf L_{m(n(k))+1,n(k)}$ of X is invariant under T; and, by part (i) of that lemma, $z \in L$. Furthermore, it follows from (13) and (11) that

$$\begin{aligned} ||Te - z|| &= \lim_{k \to \infty} ||Te - Tx_{n(k)}|| \\ &\leqslant \limsup_{k \to \infty} ||T||\ ||e - x_{n(k)}|| \\ &\leqslant \alpha ||T|| \\ &< ||Te||. \end{aligned}$$

Thus $z \neq 0$, and so $L \neq \{0\}$; it remains to prove that $L \neq X$.

Let u_k be a vector in $L_{m(n(k))+1,n(k)}$ such that $d(u_k, L_{m(n(k)),n(k)}) = 1$. If we choose a suitable vector v_k in $L_{m(n(k)),n(k)}$, and replace u_k by $u_k - v_k$, we may assume that

$$(14)\qquad \begin{cases} u_k \in L_{m(n(k))+1,n(k)}, \\ d(u_k, L_{m(n(k)),n(k)}) = 1, \qquad \|u_k\| < 2. \end{cases}$$

It follows from (8) and (9) that every element of $L_{m(n(k))+1,n(k)}$ can be expressed (uniquely) in the form $\alpha u_k + x$, where α is a complex number and $x \in L_{m(n(k)),n(k)}$. By (14),

$$(15)\qquad \|\alpha u_k + x\| \geqslant d(\alpha u_k, L_{m(n(k)),n(k)}) = |\alpha|$$

for each scalar α and each x in $L_{m(n(k)),n(k)}$.

Suppose that $L = X$; we shall obtain a contradiction. Since e, $Te \in L = \liminf L_{m(n(k))+1,n(k)}$, there exist sequences (a_k) and (b_k) such that

$$(16)\qquad a_k,\ b_k \in L_{m(n(k))+1,n(k)}$$

and

$$\|a_k - e\| \to 0, \qquad \|b_k - Te\| \to 0$$

as $k \to \infty$. We can choose scalars α_k, β_k and vectors y_k, z_k such that

$$(17)\qquad \begin{cases} y_k,\ z_k \in L_{m(n(k)),n(k)}, \\ a_k = \alpha_k u_k + y_k, \qquad b_k = \beta_k u_k + z_k. \end{cases}$$

Since the sequences (a_k) and (b_k) are bounded, it follows from (15), (17) and (14) that the same is true of (α_k), (β_k), hence also of (y_k) and (z_k). By Lemma 4.2.4, the sequences (Ty_k) and (Tz_k) converge to 0, and so

$$(18)\qquad Te = \lim Ta_k = \lim (\alpha_k Tu_k + Ty_k) = \lim \alpha_k Tu_k,$$

and similarly $T^2 e = \lim \beta_k Tu_k$. By (18) and (14),

$$0 < \|Te\| = \lim |\alpha_k|\ \|Tu_k\| \leqslant 2\|T\| \liminf |\alpha_k|,$$

so $\liminf |\alpha_k| > 0$. It follows that the sequence $(\beta_k \alpha_k^{-1})$ is ultimately bounded, and so has a subsequence converging to a scalar λ. As $k \to \infty$ through the appropriate values, we have

$$T^2 e = \lim \beta_k Tu_k = \lim (\beta_k \alpha_k^{-1}) \alpha_k Tu_k = \lambda Te,$$

contradicting (2). Hence $L \neq X$, and the proof of Theorem 4.2.1 is complete.

4.3. *Superdiagonal representations of compact linear operators*

Throughout this section, T denotes a compact linear operator acting on a complex Banach space X. The set $\mathfrak{L}$ of all closed subspaces of X is partially ordered by the inclusion relation $\subseteq$, and a totally ordered subset $\mathcal{F}$ of $\mathfrak{L}$ will be called a *chain* (more precisely, a chain of closed subspaces of X). If each subspace in a chain $\mathcal{F}$ is invariant under T, we shall describe $\mathcal{F}$ as an *invariant* chain. A trivial example of an invariant chain is the family consisting of the two subspaces $\{0\}$, X; the existence of non-trivial invariant chains can be deduced from Theorem 4.2.1.

The class $\mathcal{C}$ of all chains is itself partially ordered by the inclusion relation on subsets of $\mathfrak{L}$. Suppose that $\mathcal{C}_0$ is a totally ordered subset of $\mathcal{C}$, and let $\mathcal{F}_0 = \cup\{\mathcal{F} : \mathcal{F} \in \mathcal{C}_0\}$. If $L_1, L_2 \in \mathcal{F}_0$, then there exist $\mathcal{F}_1$, $\mathcal{F}_2$ in $\mathcal{C}_0$ such that $L_1 \in \mathcal{F}_1$ and $L_2 \in \mathcal{F}_2$. Since $\mathcal{C}_0$ is totally ordered, we may suppose that $\mathcal{F}_1 \subseteq \mathcal{F}_2$; hence, $L_1, L_2 \in \mathcal{F}_2$ and, since $\mathcal{F}_2$ is totally ordered, one of the subspaces L_1, L_2 contains the other. Thus $\mathcal{F}_0$ is a chain, and is clearly the

least upper bound of $\mathcal{C}_0$ in $\mathcal{C}$. Since each totally ordered subset of $\mathcal{C}$ has a least upper bound, it follows from Zorn's Lemma that $\mathcal{C}$ contains maximal elements, which we call *maximal chains.*

If $\mathcal{G}$ is a chain and $\mathcal{C}(\mathcal{G})$ denotes the set of all chains which contain $\mathcal{G}$, the the argument of the preceding paragraph can be applied, with $\mathcal{C}(\mathcal{G})$ in place of $\mathcal{C}$, to prove that $\mathcal{C}(\mathcal{G})$ contains a maximal element $\mathcal{G}_1$. In fact, $\mathcal{G}_1$ is maximal in the class $\mathcal{C}$ of *all* chains; for if $\mathcal{F} \in \mathcal{C}$ and $\mathcal{G}_1 \subseteq \mathcal{F}$ then $\mathcal{G} \subseteq \mathcal{G}_1 \subseteq \mathcal{F}$; thus $\mathcal{F} \in \mathcal{C}(\mathcal{G})$, and the maximality of $\mathcal{G}_1$ in $\mathcal{C}(\mathcal{G})$ implies that $\mathcal{F} = \mathcal{G}_1$. This shows that every chain $\mathcal{G}$ is contained in at least one maximal chain $\mathcal{G}_1$.

Similar reasoning shows that the class $\mathcal{C}_i$ of all invariant chains contains maximal elements, which we call *maximal invariant chains,* and that each invariant chain is contained in at least one maximal invariant chain. It is not immediately obvious that maximal elements of $\mathcal{C}_i$ $(\subseteq \mathcal{C})$ are maximal in $\mathcal{C}$; that is, it is not at once apparent that maximal invariant chains are necessarily maximal chains. However, we prove in Lemma 4.3.3 that this is so.

The norm closure of a subset S of X will be denoted by $c\ell(S)$. Given a subfamily $\mathcal{F}_0$ of a chain $\mathcal{F}$, the set $\cap\{L : L \in \mathcal{F}_0\}$ is a closed subspace of X; the same is true of $c\ell[\cup\{L : L \in \mathcal{F}_0\}]$, since $\mathcal{F}_0$ is totally ordered by inclusion. Given M in $\mathcal{F}$, we define a closed subspace M_- of X by

(1) $$M_- = c\ell[\cup\{L : L \in \mathcal{F},\ L \subsetneqq M\}],$$

interpreting the right-hand side as $\{0\}$ if there is no subspace L in $\mathcal{F}$ which is properly contained in M. Of course, M_- depends on the chain $\mathcal{F}$, as well as on M. We describe $\mathcal{F}$ as a *simple chain* if it satisfies the following conditions

(i) $\{0\} \in \mathcal{F}, \quad X \in \mathcal{F}$;

(ii) if $\mathcal{F}_0$ is a subfamily of $\mathcal{F}$, then the closed subspaces
$$\cap\{L : L \in \mathcal{F}_0\}, \qquad c\ell[\cup\{L : L \in \mathcal{F}_0\}]$$
are in $\mathcal{F}$;

(iii) for each M in $\mathcal{F}$, the quotient space M/M_- is at most 1-dimensional.

Condition (ii) implies that $M_- \in \mathcal{F}$, for each M in $\mathcal{F}$; if M has an immediate predecessor L in $\mathcal{F}$, then $M_- = L$, but $M_- = M$ when M has no immediate predecessor. By a *continuous chain* we mean a simple chain $\mathcal{F}$ such that $M = M_-$ for each M in $\mathcal{F}$.

LEMMA 4.3.1. *Each simple chain is maximal.*

Proof. Suppose the contrary. Then there is a simple chain $\mathcal{F}$, and a chain $\mathcal{G}$ which properly contains $\mathcal{F}$. If N is a subspace in $\mathcal{G} \sim \mathcal{F}$ then $N \neq \{0\}$, X (since $\{0\}, X \in \mathcal{F}$). Let

$$M = \cap\{L : L \in \mathcal{F},\ N \subseteq L\},$$
$$M' = c\ell[\cup\{L : L \in \mathcal{F},\ L \subseteq N\}].$$

Since $\mathcal{F}$ is simple, it follows that $M, M' \in \mathcal{F}$; it is clear that $M' \subseteq N \subseteq M$ and, since $N \notin \mathcal{F}$,

(2) $$M' \subsetneqq N \subsetneqq M.$$

We now show that $M' = M_-$. Since $M' \in \mathcal{F}$ and $M' \subsetneqq M$, the inclusion $M' \subseteq M_-$ is an immediate consequence of (1). To prove the reverse inclusion it is sufficient, in view of (1), to show that M' contains each subspace L in $\mathcal{F}$ for which $L \subsetneqq M$. Given such a subspace L, it follows from the definition of M that $N \not\subseteq L$. Since $L, N \in \mathcal{G}$ (which is totally ordered) we deduce that $L \subseteq N$, and so $L \subseteq M'$ by the definition of M'. This proves that $M' = M_-$, and (2) implies that

the quotient space M/M_- has dimension at least 2, contradicting our assumption that $\mathfrak{F}$ is a simple chain.

LEMMA 4.3.2. *Each maximal invariant chain is simple.*

Proof. Suppose that $\mathfrak{F}$ is a maximal invariant chain. Then $\{0\}$, $X \in \mathfrak{F}$, since otherwise $\mathfrak{F}$ could be enlarged by the addition of these subspaces. If $\mathfrak{F}_0$ is a subfamily of $\mathfrak{F}$, consider the closed subspace

$$N = \cap\{L : L \in \mathfrak{F}_0\}$$

of X. Since each L in $\mathfrak{F}_0$ is invariant under T, the same is true of N. Given any subspace M in $\mathfrak{F}$, the total ordering of $\mathfrak{F}$ implies that *either* (a) $M \subseteq L$ for each L in $\mathfrak{F}_0$, and so $M \subseteq N$, *or* (b) $L \subseteq M$ for at least one L in $\mathfrak{F}_0$, and so $N \subseteq M$. It follows that $\mathfrak{F} \cup \{N\}$ is totally ordered by inclusion and is therefore an invariant chain; by maximality of $\mathfrak{F}$, $N \in \mathfrak{F}$. A similar argument shows that the closed subspace

$$c\ell[\cup\{L : L \in \mathfrak{F}_0\}]$$

is in $\mathfrak{F}$.

So far, we have shown that $\mathfrak{F}$ has the first two of the defining properties of simple chains; it remains to verify the third. Suppose that the quotient space M/M_- has dimension greater than 1, for some M in $\mathfrak{F}$. Since M and M_- are invariant under the compact linear operator T, it follows from Theorem 1.8.3(i) and (ii) that the mapping $T_0 : x+M_- \to Tx+M_-$ $(x \in M)$ is a compact linear operator acting on the Banach space M/M_-. By Theorem 4.2.1 there is a closed subspace N_0 of M/M_- such that $\{0\} \neq N_0 \neq M/M_-$ and $T_0(N_0) \subseteq N_0$. It follows easily from these conditions that, if

$$N = \{x \in M : x+M_- \in N_0\},$$

then N is a closed subspace of X and

$$(3) \qquad M_- \subsetneq N \subsetneq M, \qquad T(N) \subseteq N.$$

Given any subspace L in $\mathfrak{F}$, we have *either* (a) $M \subseteq L$, and so $N \subsetneq L$, *or* (b) $L \subsetneq M$ and, by (1), $L \subseteq M_- \subsetneq N$. Hence $N \notin \mathfrak{F}$ and $\mathfrak{F} \cup \{N\}$ is totally ordered by inclusion. This, together with (3), shows that $\mathfrak{F} \cup \{N\}$ is an invariant chain which properly contains $\mathfrak{F}$, contradicting the maximality assumption. It follows that, for each M in $\mathfrak{F}$, the quotient space M/M_- has dimensions at most 1. This completes the proof that $\mathfrak{F}$ is a simple chain.

LEMMA 4.3.3. *A chain $\mathfrak{F}$ is maximal if and only if it is simple. For an invariant chain $\mathfrak{F}$, the following three conditions are equivalent:*

(i) $\mathfrak{F}$ is a maximal chain;

(ii) $\mathfrak{F}$ is a maximal invariant chain;

(iii) $\mathfrak{F}$ is simple.

Proof. Since a maximal chain is a maximal invariant chain for the compact linear operator 0, it follows from Lemma 4.3.2 that maximal chains are simple (of course, Theorem 4.2.1 is not needed to prove Lemma 4.3.2 in the case $T = 0$). This, together with Lemma 4.3.1, shows that a chain $\mathfrak{F}$ is maximal if and only if it is simple.

Suppose now that $\mathfrak{F}$ is an invariant chain for a compact linear operator T. If $\mathfrak{F}$ is maximal in the class of *all* chains, it is *a fortiori* maximal in the class of invariant chains, so (i) implies (ii). It has already been noted, in the earlier lemmas, that (ii) implies (iii) and (iii) implies (i).

THEOREM 4.3.4. *Let T be a compact linear operator acting on a complex Banach space X. Then there is a simple chain $\mathcal{F}$ of closed subspaces of X such that each L in $\mathcal{F}$ is invariant under T.*

Proof. In view of Lemma 4.3.2 it is sufficient to take for $\mathcal{F}$ any maximal invariant chain.

Throughout the remainder of §4.3, we shall use the symbols T, X, $\mathcal{F}$ with the meanings attributed to them in Theorem 4.3.4. If $M \in \mathcal{F}$ then either $M = M_-$ or M/M_- has dimension 1. In the latter case, suppose that

$$z_M \in M \sim M_-, \tag{4}$$

so that M is the linear span of $\{z_M\} \cup M_-$. Since M is invariant under T, $Tz_M \in M$; thus there exist a scalar α_M and a vector y_M such that

$$Tz_M = \alpha_M z_M + y_M, \qquad y_M \in M_-. \tag{5}$$

It is easily verified that α_M does not depend on the particular choice of z_M in $M \sim M_-$. Since M_- is invariant under $T-\alpha_M I$, and $(T-\alpha_M I)z_M \in M_-$, it follows that

$$(T-\alpha_M I)(M) \subseteq M_-. \tag{6}$$

When $M \in \mathcal{F}$ and $M = M_-$, we define $\alpha_M = 0$. In this way we associate with each M in $\mathcal{F}$ a scalar α_M which we call the *diagonal coefficient* of T at M. In the finite-dimensional case, the elements $\{z_M : M \in \mathcal{F}\}$ form a basis, which inherits a total ordering from $\mathcal{F}$; the matrix of T with respect to this ordered basis is superdiagonal, and has diagonal entries α_M $(M \in \mathcal{F})$.

The *diagonal multiplicity* of a scalar α is defined to be the (possibly infinite) cardinal number of the set $\{M : M \in \mathcal{F},\ \alpha_M = \alpha\}$.

LEMMA 4.3.5. *Suppose that K is a compact linear operator acting on a complex Banach space X, and δ is a positive real number. Then there is a finite subset S of the unit ball $X_1 = \{x \in X : \|x\| \leqslant 1\}$ of X which has the following property: if $x \in X_1$, there exists y in S such that $\|Kx-Ky\| < \delta$.*

Proof. Suppose that no such set S exists. We construct an infinite sequence (x_n) of elements of X_1, inductively, as follows: x_1 is an arbitrary vector in X_1 and, when $x_1, \ldots, x_{n-1}$ have been selected, we choose x_n in X_1 so that $\|Kx_n-Kx_m\| \geqslant \delta$ for $1 \leqslant m < n$. The existence of a suitable vector x_n follows from the assumption that the finite subset $\{x_1, \ldots, x_{n-1}\}$ of X_1 cannot have the property required for S in the lemma.

In this way we obtain a bounded infinite sequence (x_n) of elements of X such that $\|Kx_m-Kx_n\| \geqslant \delta$ whenever $m \neq n$. Clearly (Kx_n) has no convergent subsequence, contradicting our assumption that K is compact. This proves the existence of a set S with the stated property.

LEMMA 4.3.6. *Given a positive real number δ and a non-zero subspace M in $\mathcal{F}$, there exists L in $\mathcal{F}$ such that $L \subsetneqq M$ and $d(Tx, L) \leqslant \delta\|x\|$ for each x in M_-.*

Remark. The interest of this lemma lies in the case in which $M = M_-$; the result is trivial, with $L = M_-$, when $M \neq M_-$.

Proof. By Lemma 4.3.5, with K the restriction of T to M_-, there is a finite subset S of the unit ball of M_- with the following property: if $x \in M_-$ and $\|x\| \leqslant 1$, there exists y in S such that $\|Tx-Ty\| < \frac{1}{2}\delta$. If $y \in S$, then $y \in M_-$ and so

$$Ty \in M_- = c\ell[\cup\{L : L \in \mathcal{F},\ L \subsetneqq M\}];$$

so there is a subspace L_y such that

$$L_y \in \mathfrak{F}, \qquad L_y \subsetneq M, \qquad d(Ty, L_y) < \tfrac{1}{2}\delta \quad (y \in S).$$

Since S is finite and $\mathfrak{F}$ is totally ordered by inclusion, the family $\{L_y : y \in S\}$ has a largest member L, and

$$L \in \mathfrak{F}, \qquad L \subsetneq M, \qquad d(Ty, L) < \tfrac{1}{2}\delta \quad (y \in S).$$

If $x \in M_-$ and $||x|| \leqslant 1$, there exists y in S such that $||Tx-Ty|| < \tfrac{1}{2}\delta$; thus

$$\begin{aligned} d(Tx, L) &\leqslant d(Tx-Ty, L)+d(Ty, L) \\ &\leqslant ||Tx-Ty||+d(Ty, L) < \delta. \end{aligned}$$

By homogeneity, $d(Tx, L) \leqslant \delta||x||$ for each x in M_-.

LEMMA 4.3.7. *Suppose that λ is a non-zero eigenvalue of T, x is a non-zero vector satisfying $Tx = \lambda x$ and*

$$M = \cap\{L : L \in \mathfrak{F},\ x \in L\}.$$

Then $M \in \mathfrak{F}$, $x \in M \sim M_-$ and $\alpha_M = \lambda$.

Proof. Since $\mathfrak{F}$ is a simple chain, $M \in \mathfrak{F}$; clearly M is the smallest subspace in $\mathfrak{F}$ which contains x, and $M \neq \{0\}$.

If $M = M_-$, let δ be a real number such that $0 < \delta < |\lambda|$, and choose L satisfying the conclusions of Lemma 4.3.6. Since $L \in \mathfrak{F}$ and $L \subsetneq M$, it follows that $x \notin L$, so $d(x, L) > 0$. Furthermore, if $y \in L$ then $Ty \in L$ and thus

$$\begin{aligned} |\lambda| d(x, L) = d(\lambda x, L) &= d(Tx, L) \\ &= d(Tx+Ty, L) \\ &\leqslant \delta||x+y||; \end{aligned}$$

the last inequality follows from Lemma 4.3.6, since $x+y \in M = M_-$. By taking lower bounds as y varies in L, we obtain $|\lambda| d(x, L) \leqslant \delta d(x, L)$, contradicting our earlier assertions that $0 < \delta < |\lambda|$ and $d(x, L) > 0$. Thus $M \neq M_-$.

Since M is the smallest subspace in $\mathfrak{F}$ which contains x, and $M_- \subsetneq M$, it follows that $x \notin M_-$. However, by (6),

$$(\lambda-\alpha_M)x = (T-\alpha_M I)x \in M_-,$$

so $\lambda-\alpha_M = 0$.

The lemma just proved shows that a non-zero eigenvalue of T is a diagonal coefficient. We now have a result in the opposite direction.

LEMMA 4.3.8. *If $M \in \mathfrak{F}$ and $\alpha_M \neq 0$, then α_M is an eigenvalue of T. If α_M has index 1 relative to T, there is a vector x in $M \sim M_-$ satisfying $Tx = \alpha_M x$.*

Proof. Let T^M be the compact linear operator obtained by restricting T to M, and let $\mathfrak{N}$ and $\mathfrak{R}$ denote the null space and (closed) range space, respectively, of $T^M-\alpha_M I$. Since $\alpha_M \neq 0$, we have $M \neq M_-$; and, by (6)

$$\mathfrak{R} = (T^M-\alpha_M I)(M) = (T-\alpha_M I)(M) \subseteq M_- \subsetneq M.$$

It follows from Theorem 1.8.1 that α_M is an eigenvalue of T_M (and hence, also, of T).

If α_M has index 1 relative to T, the operators $T-\alpha_M I$ and $(T-\alpha_M I)^2$ have the same null space; this property is preserved when we restrict to M, so α_M has index 1 as an eigenvalue of T_M. By Theorem 1.8.1,

$$M = \mathfrak{R}+\mathfrak{N}.$$

Since $\mathcal{R} \subseteq M_-$, it follows that $\mathcal{N}$ meets $M \sim M_-$; that is, there exists x in $M \sim M_-$ such that $\alpha_M x = T^M x = Tx$.

LEMMA 4.3.9. *If λ is a non-zero diagonal coefficient of T, then its diagonal multiplicity is finite and is equal to its algebraic multiplicity as an eigenvalue of T.*

Proof. Let d denote the diagonal multiplicity, m the algebraic multiplicity and ν the index of λ relative to T. For each integer $k(\geqslant 0)$, let $\mathcal{N}_k$ be the null space of $(T-\lambda I)^k$. Then $\mathcal{N}_\nu$ is m-dimensional, and

$$\{0\} = \mathcal{N}_0 \subsetneq \mathcal{N}_1 \subsetneq \cdots \subsetneq \mathcal{N}_\nu = \mathcal{N}_{\nu+1} = \mathcal{N}_{\nu+2} = \cdots$$

We show first that it is sufficient to consider the case in which $\nu = 1$. For this purpose we introduce the compact linear operator S defined by

$$\mu I - S = (\lambda I - T)^\nu,$$

where $\mu = \lambda^\nu$. Since $\mu I - S$ and $(\mu I - S)^2$ have the same null space $\mathcal{N}_\nu (= \mathcal{N}_{2\nu})$, of dimension m, it follows that μ is an eigenvalue of S with index 1 and algebraic multiplicity m. Since S is a polynomial in T, each subspace in the simple chain $\mathcal{F}$ is invariant under S, and we may consider the diagonal coefficients β_M $(M \in \mathcal{F})$ of S. If $M \in \mathcal{F}$, $M \neq M_-$ and z_M, y_M satisfy (4) and (5), then

$$\begin{aligned} Sz_M &= \mu z_M - (\mu I - S) z_M \\ &= \mu z_M - (\lambda I - T)^\nu z_M \\ &= \mu z_M - [(\lambda - \alpha_M) I + (\alpha_M I - T)]^\nu z_M \\ &= [\mu - (\lambda - \alpha_M)^\nu] z_M + \sum_{r=1}^{\nu} \binom{\nu}{r} (\lambda - \alpha_M)^{\nu - r} (\alpha_M I - T)^r z_M \\ &= [\mu - (\lambda - \alpha_M)^\nu] z_M - v_M, \end{aligned}$$

where, by (6), $v_M \in M_-$. Hence

$$\beta_M = \mu - (\lambda - \alpha_M)^\nu \tag{7}$$

(the same equation is satisfied when $M \in \mathcal{F}$ and $M = M_-$, since $\alpha_M = \beta_M = 0$ in that case, and $\mu = \lambda^\nu$). It follows that $\beta_M = \mu$ if and only if $\alpha_M = \lambda$, so μ has diagonal multiplicity d as a diagonal coefficient of S. It is now sufficient to prove the lemma under the additional hypothesis that λ has index 1 relative to T, since in the general case we can reduce to this situation by considering S, μ in place of T, λ respectively.

Suppose, therefore, that $\nu = 1$. If $d > m$, there exist subspaces $M(0), M(1), \ldots, M(m)$ in $\mathcal{F}$ such that

$$M(0) \subsetneq M(1) \subsetneq \cdots \subsetneq M(m)$$

and $\alpha_{M(j)} = \lambda$ $(j = 0, \ldots, m)$. Since $M(j-1) \subsetneq M(j)$, it follows that

$$M(j-1) \subseteq M(j)_- \quad (j = 1, \ldots, m).$$

By Lemma 4.3.8 there exist vectors $x_0, x_1, \ldots, x_m$ such that $Tx_j = \lambda x_j$ and

$$x_j \in M(j) \sim M(j)_- \quad (j = 0, 1, \ldots, m). \tag{8}$$

The vectors $x_0, x_1, \ldots, x_m$ lie in the (m-dimensional) null space of $T - \lambda I$, and are therefore linearly dependent. Hence some x_j $(1 \leqslant j \leqslant m)$ is a linear combination of its predecessors and, for suitable scalars $\lambda_0, \lambda_1, \ldots, \lambda_{j-1}$,

$$x_j = \lambda_0 x_0 + \lambda_1 x_1 + \cdots + \lambda_{j-1} x_{j-1} \in M(j-1) \subseteq M(j)_-,$$

contradicting (8). Thus $d \leqslant m$, and so d is finite.

There are just d distinct subspaces M in $\mathcal{F}$ such that $\alpha_M = \lambda$; let these be $M(1), \ldots, M(d)$. Since $\alpha_{M(j)} \neq 0$, the quotient space $M(j)/M(j)_-$ has dimension 1, and there is a continuous linear functional ψ_j on $M(j)$, with kernel $\psi_j^{-1}(0)$ equal to $M(j)_-$. By the

Hahn–Banach theorem, ψ_j extends to a continuous linear functional φ_j on X, and we have

$$M(j)_- = \{x \in M(j) : \varphi_j(x) = 0\}. \tag{9}$$

If $m > d$, there is a non-zero vector x_0 in the (m-dimensional) null space of $T-\lambda I$ which satisfies the d linear conditions $\varphi_j(x_0) = 0$ ($j = 1, \ldots, d$). If $M = \cap\{L : L \in \mathfrak{F},\ x_0 \in L\}$, then Lemma 4.3.7 asserts that $M \in \mathfrak{F}$, $x_0 \in M \sim M_-$ and $\alpha_M = \lambda$. The last condition implies that $M = M(j)$ for some j ($1 \leqslant j \leqslant d$), and we have $x_0 \in M(j) \sim M(j)_-$, $\varphi_j(x_0) = 0$, contradicting (9). It follows that $m \leqslant d$; since we have already proved the reverse inequality, $m = d$.

The following theorem summarises the main results obtained in the preceding lemmas.

THEOREM 4.3.10. *Suppose that T is a compact linear operator acting on a complex Banach space X, and $\mathfrak{F}$ is a simple chain of closed subspaces of X, each of which is invariant under T. Then*

(i) a non-zero scalar λ is an eigenvalue of T if and only if it is a diagonal coefficient of T;

(ii) the diagonal multiplicity of λ is equal to its algebraic multiplicity as an eigenvalue of T;

(iii) the operator T is quasi-nilpotent if and only if $T(M) \subseteq M_-$ for each M in $\mathfrak{F}$.

Proof. The only statement not already proved is (iii). Now the spectrum of T consists of 0 (except, possibly, in the finite-dimensional case) and the non-zero eigenvalues of T. From part (i) of the theorem it follows that T is quasi-nilpotent if and only if $\alpha_M = 0$ whenever $M \in \mathfrak{F}$. If $M \in \mathfrak{F}$ and $M \neq M_-$, then (4), (5) and (6) imply that $\alpha_M = 0$ if and only if $T(M) \subseteq M_-$; when $M = M_-$, the conditions $\alpha_M = 0$ and $T(M) \subseteq M_-$ are both satisfied automatically. Hence T is quasi-nilpotent if and only if $T(M) \subseteq M_-$ for each M in $\mathfrak{F}$.

COROLLARY 4.3.11. *If T is a compact linear operator acting on a complex Banach space X, and there is a continuous chain $\mathfrak{F}$ of closed subspaces of X, each of which is invariant under T, then T is quasi-nilpotent.*

Proof. For each M in $\mathfrak{F}$, $T(M) \subseteq M = M_-$, so the result follows from part (iii) of Theorem 4.3.10.

EXAMPLE 4.3.12. Suppose that X is the Hilbert space $L_2(0,1)$ (Lebesgue measure) and that, when $0 < \lambda < 1$, L_λ is the closed subspace of X defined by

$$L_\lambda = \{x \in L_2(0,1) : x(s) = 0 \text{ almost everywhere on } [0,\lambda]\}.$$

Since $L_\lambda \subsetneq L_\mu$ when $0 \leqslant \mu < \lambda \leqslant 1$, the family $\mathfrak{F} = \{L_\lambda : 0 \leqslant \lambda \leqslant 1\}$ is a chain of closed subspaces of X. It is clear that

$$L_1 = \{0\}, \qquad L_0 = X. \tag{10}$$

We assert also that, if S is a non-empty subset of $[0,1]$ and $\alpha = \inf S$, $\beta = \sup S$, then

$$\cap\{L_\lambda : \lambda \in S\} = L_\beta, \qquad c\ell[\cup\{L_\lambda : \lambda \in S\}] = L_\alpha. \tag{11}$$

As a special case of the second relation, we have

$$c\ell[\cup\{L_\lambda : \alpha < \lambda \leqslant 1\}] = L_\alpha \quad (0 \leqslant \alpha < 1). \tag{12}$$

To prove the first equality in (11), let $(\lambda(n))$ be an increasing sequence of elements of S, with the limit β. If

$$x \in \cap\{L_\lambda : \lambda \in S\},$$

then $x \in L_{\lambda(n)}$ and so $x(s) = 0$ for almost all s in $[0,\lambda(n)]$ $(n = 1, 2, \ldots)$; thus $x(s) = 0$ for almost all s in $[0,\beta]$, and $x \in L_\beta$. This shows that

$$\cap\{L_\lambda : \lambda \in S\} \subseteq L_\beta,$$

and the reverse inclusion is apparent. To prove the second part of (11), suppose that $x \in L_\alpha$, and let $(\mu(n))$ be a decreasing sequence of elements of S, with limit α. If x_n is the element of $L_2(0,1)$ which vanishes on $[0,\mu(n)]$ and coincides with x on $[\mu(n),1]$, then $x_n \in L_{\mu(n)}$ and, since x vanishes on $[0,\alpha]$,

$$\begin{aligned}||x-x_n||^2 &= \int_0^1 |x(s)-x_n(s)|^2 ds\\ &= \int_\alpha^{\mu(n)} |x(s)|^2 ds \to 0\end{aligned}$$

as $n \to \infty$. Thus

$$x \in c\ell[\cup\{L_{\mu(n)} : n = 1, 2, \ldots\}] \subseteq c\ell[\cup\{L_\lambda : \lambda \in S\}].$$

It follows that $L_\alpha \subseteq c\ell[\cup\{L_\lambda : \lambda \in S\}]$; since the reverse inclusion is obvious, the proof of (11) is complete.

Since every subfamily of $\mathfrak{F}$ has the form $\{L_\lambda : \lambda \in S\}$ for some subset S of $[0,1]$, we can deduce from (10), (11) and (12) that $\mathfrak{F}$ has the three defining properties of simple chains (with $M_- = M$ for each M in $\mathfrak{F}$). Thus $\mathfrak{F}$ is a continuous chain of closed subspaces of X.

Let h be a Lebesgue measurable complex-valued function on the square $[0,1] \times [0,1]$, and suppose that

$$h(s,\ t) = 0 \quad (0 \leqslant s < t \leqslant 1),$$

$$\int_0^1 ds \int_0^s |h(s,\ t)|^2 dt < \infty.$$

The associated integral operator T_h acting on $L_2(0,1)$ is defined by

$$(T_h x)(s) = \int_0^s h(s,\ t)x(t)dt \qquad (13)$$

$$(x \in L_2(0,1) \ :\ 0 \leqslant s \leqslant 1).$$

If $0 \leqslant \lambda \leqslant 1$ and $x \in L_\lambda$, then the function x vanishes almost everywhere on $[0,\lambda]$; it follows from (13) that $T_h x$ has the same property, so $T_h x \in L_\lambda$. Thus T_h leaves invariant each subspace in the continuous chain $\mathfrak{F}$ and is therefore quasi-nilpotent, by Corollary 4.3.11. Operators of the type defined in (13), in which $(T_h x)(s)$ is an integral over $[0,s]$, are known as *Volterra integral operators*.

It follows easily from the preceding discussion that, if $\mathcal{H}$ is a separable infinite-dimensional Hilbert space, then there is a continuous chain of closed subspaces of $\mathcal{H}$. To prove this, note first that $L_2(0,1)$ is infinite-dimensional, and is separable since polynomials, in which each coefficient has rational real and imaginary parts, form a countable everywhere dense subset. Thus $\mathcal{H}$ and $L_2(0,1)$ are both isometrically isomorphic to $\ell_2(N)$, where N is the set of positive integers. It follows that there is an isometric isomorphism U from $L_2(0,1)$ onto $\mathcal{H}$; and it is apparent that, if $\mathfrak{F}$ is defined as above, then the family $\{U(L) : L \in \mathfrak{F}\}$ is a continuous chain of subspaces of $\mathcal{H}$.

4.4. *Superdiagonal integrals*

This section is concerned with the superdiagonal representation of a compact linear operator acting on a Hilbert space. In this context, the theory developed in §4.3 can be strengthened considerably by use of the one-to-one order preserving correspondence between closed subspaces and projections (see §1.7, especially Lemma 1.7.10 and Theorem 1.7.11). Our main interest is in the

following question: given a compact self-adjoint operator K acting on a Hilbert space $\mathcal{H}$, and a simple chain $\mathcal{F}$ of closed subspaces of $\mathcal{H}$, in what circumstances is there a quasi-nilpotent operator T which has skew-adjoint part K and leaves invariant each subspace in $\mathcal{F}$? Before considering this problem, we shall express some of the results of §4.3 in a slightly different form.

The projection from a Hilbert space $\mathcal{H}$ onto a closed subspace M will be denoted by $E(M)$. If $T \in \mathcal{B}(\mathcal{H})$ and M is invariant under T then, given any x in $\mathcal{H}$, we have $TE(M)x \in T(M) \subseteq M$ and therefore $TE(M)x = E(M)TE(M)x$; thus $TE(M) = E(M)TE(M)$. Conversely, if $T \in \mathcal{B}(\mathcal{H})$ and $TE(M) = E(M)TE(M)$ then, for each x in M,

$$Tx = TE(M)x = E(M)TE(M)x \in M;$$

and thus, $T(M) \subseteq M$. It follows that a closed subspace M of $\mathcal{H}$ is invariant under T if and only if $TE(M) = E(M)TE(M)$.

Suppose now that T is a compact linear operator acting on a Hilbert space $\mathcal{H}$, and $\mathcal{F}$ is a simple chain of closed subspaces of $\mathcal{H}$, each of which is invariant under T. Then

$$(1) \qquad TE(M) = E(M)TE(M) \quad (M \in \mathcal{F})$$

and, since $\mathcal{F}$ is totally ordered by inclusion,

$$(2) \qquad E(L)E(M) = E(M)E(L) = E(L \wedge M) \quad (L,\ M \in \mathcal{F}).$$

Note also that, if $L,\ M \in \mathcal{F}$, and $L \subseteq M$, then $E(M)-E(L) = E(M \cap L^{-})$. Equation 4.3(1) implies that

$$(3) \qquad E(M_{-}) = \vee\{E(L) : L \in \mathcal{F},\ L \subsetneq M\} \quad (M \in \mathcal{F}).$$

If $M \in \mathcal{F}$ and $M \neq M_{-}$, the element z_M of 4.3(4) can be chosen so that

$$(4) \qquad z_M \in M \cap (M_{-})^{\perp}, \qquad \|z_M\| = 1 \quad (M \in \mathcal{F},\ M \neq M_{-}).$$

It then follows from 4.3(5) that

$$\langle Tz_M, z_M\rangle = a_M\langle z_M, z_M\rangle + \langle y_M, z_M\rangle = a_M,$$

since $y_M \in M_{-}$; so the diagonal coefficient a_M of T at M is given by

$$(5) \qquad a_M = \langle Tz_M, z_M\rangle \quad (M \in \mathcal{F},\ M \neq M_{-}).$$

For each M in $\mathcal{F}$, the projection $E(M)-E(M_{-})$ will be denoted by $P(M)$. If $M \neq M_{-}$, then M is the linear span of z_M and M_{-} and so, for each x in $\mathcal{H}$, there exist a scalar a and a vector y in M_{-} such that $E(M)x = az_M + y$. It follows that

$$E(M_{-})x = E(M_{-})E(M)x = y,$$

$$P(M)x = E(M)x - E(M_{-})x = az_M.$$

Since $y \in M_{-}$ and $z_M \in (M_{-})^{\perp}$,

$$a = \langle az_M + y, z_M\rangle = \langle E(M)x, z_M\rangle$$
$$= \langle x, E(M)z_M\rangle = \langle x, z_M\rangle.$$

Thus

$$(6) \qquad P(M) = E(M)-E(M_{-}), \qquad P(M)x = \langle x, z_M\rangle z_M$$

whenever $M \in \mathcal{F}$, $M \neq M_{-}$ and $x \in \mathcal{H}$. A straightforward computation, based on (6) and (5), now gives

$$(7) \qquad P(M)TP(M) = a_M P(M) \quad (M \in \mathcal{F},\ M \neq M_{-}).$$

The following result, which is a variant of Theorem 4.3.10(iii), is an immediate consequence of part (i) of that theorem, together with (7).

LEMMA 4.4.1. *Suppose that T is a compact linear operator acting on a Hilbert space $\mathcal{H}$ and $\mathcal{F}$ is a simple chain of closed*

subspaces of $\mathcal{H}$*, each of which is invariant under* T*. Then* T *is quasi-nilpotent if and only if* $P(M)TP(M) = 0$ *whenever* $M \in \mathcal{F}$ *and* $M \neq M_-$.

We now turn to the problem of main interest in this section. Our first result is superficial, but is sometimes useful.

THEOREM 4.4.2. *Suppose that* $\mathcal{F}$ *is a simple chain of closed subspaces of a Hilbert space* $\mathcal{H}$*, and* K *is a compact self-adjoint operator acting on* $\mathcal{H}$*. A necessary condition for the existence of a quasi-nilpotent operator* T *which has skew-adjoint part* K *and leaves invariant each subspace in* $\mathcal{F}$ *is*

$$P(M)KP(M) = 0 \quad (M \in \mathcal{F},\ M \neq M_-). \tag{8}$$

If such an operator T *exists, it is unique, and is a compact linear operator.*

Proof. Suppose that there is an operator T with the stated properties. Since $\operatorname{Im}(T)$ $(= K)$ is compact, it follows from Theorem 1.8.8. that T is compact, and Lemma 4.4.1 asserts that $P(M)TP(M) = 0$ whenever $M \in \mathcal{F}$ and $M \neq M_-$. By taking adjoints, we obtain $P(M)T^*P(M) = 0$, so

$$P(M)KP(M) = \tfrac{1}{2}iP(M)(T^*-T)P(M) = 0 \quad (M \in \mathcal{F},\ M \neq M_-).$$

If T_1 and T_2 are two such operators, the above reasoning shows that both are compact and satisfy $P(M)TP(M)=0$ $(M \in \mathcal{F},\ M \neq M_-)$. Thus T_1-T_2 is compact, leaves invariant each subspace in $\mathcal{F}$, and satisfies $P(M)(T_1-T_2)P(M) = 0$ $(M \in \mathcal{F},\ M \neq M_-)$. By Lemma 4.4.1, T_1-T_2 is quasi-nilpotent; it is also self-adjoint, since

$$\operatorname{Im}(T_1-T_2) = \operatorname{Im} T_1 - \operatorname{Im} T_2 = K-K = 0,$$

so it follows from Corollary 1.7.5 that $T_1-T_2 = 0$.

Our next objective is to find an explicit formula for T, in terms of $\mathcal{F}$ and K, when T, $\mathcal{F}$, K are related as in Theorem 4.4.2. Before this can be done, we need some auxiliary results.

LEMMA 4.4.3. *Suppose that* $\mathcal{F}$ *is a simple chain of closed subspaces of a Hilbert space* $\mathcal{H}$*,* K *is a compact linear operator acting on* $\mathcal{H}$*,* $P(M)KP(M) = 0$ *whenever* $M \in \mathcal{F}$ *and* $M \neq M_-$*, and* δ *is a positive real number.*

(i) If $M \in \mathcal{F}$ *and* $M \neq \{0\}$*, there exists* L *in* $\mathcal{F}$ *such that* $L \subsetneqq M$ *and*

$$\|[E(M)-E(L)]K[E(M)-E(L)]\| \leqslant \delta.$$

(ii) If $M \in \mathcal{F}$ *and* $M \neq \mathcal{H}$*, there exists* N *in* $\mathcal{F}$ *such that* $M \subsetneqq N$ *and*

$$\|[E(N)-E(M)]K[E(N)-E(M)]\| \leqslant \delta.$$

Proof. By Lemma 4.3.5 there is a finite subset S of the unit ball $\mathcal{H}_1 = \{x \in \mathcal{H} : \|x\| \leqslant 1\}$ of $\mathcal{H}$ which has the following property: given any x in $\mathcal{H}_1$, there exists y in S such that $\|Kx-Ky\| < \tfrac{1}{2}\delta$. The set

$$U = \{A \in \mathcal{B}(\mathcal{H}) : \|AKy-E(M)Ky\| < \tfrac{1}{2}\delta \text{ for each } y \text{ in } S\}$$

is a neighbourhood of $E(M)$ in the strong operator topology.

We now prove (i). If $M \neq M_-$, then

$$[E(M)-E(M_-)]K[E(M)-E(M_-)] = P(M)KP(M) = 0,$$

and it suffices to take $L = M_-$. Now suppose that $M = M_-$, so that

$$E(M) = E(M_-) = \vee\{E(L) : L \in \mathcal{F},\ L \subsetneqq M\}.$$

By Theorem 1.7.11, the set $\{E(L) : L \in \mathcal{F},\ L \subsetneqq M\}$ has $E(M)$ in its strong closure, and therefore meets the strong neighbourhood U of $E(M)$. Hence there is a subspace L satisfying

(9) $$L \in \mathcal{F},\ L \subsetneq M,\ ||E(L)Ky - E(M)Ky|| < \tfrac{1}{2}\delta \quad (y \in S).$$

Given any x in $\mathcal{H}_1$, there exists y in S such that $||Kx-Ky|| < \tfrac{1}{2}\delta$. This, with (9), gives

$$\begin{aligned} &||[E(M)-E(L)]Kx|| \\ &\qquad \leqslant ||[E(M)-E(L)](Kx-Ky)||+||[E(M)-E(L)]Ky|| \\ &\qquad < ||Kx-Ky||+\tfrac{1}{2}\delta < \delta; \end{aligned}$$

so $||[E(M)-E(L)]Kx|| < \delta$ whenever $x \in \mathcal{H}_1$. Thus

$$||[E(M)-E(L)]K|| \leqslant \delta,$$

and this implies the inequality required in (i).

To prove (ii), we consider the subspace

(10) $$L = \cap\{N : N \in \mathcal{F},\ M \subsetneq N\}.$$

Then $M \subset L$ and, since $\mathcal{F}$ is a simple chain, $L \in \mathcal{F}$; clearly there is no subspace N in $\mathcal{F}$ such that $M \subsetneq N \subsetneq L$. If $L \neq M$, then M is the immediate predecessor of L in $\mathcal{F}$, so $M = L_-$; in this case

$$[E(L)-E(M)]K[E(L)-E(M)] = P(L)KP(L) = 0,$$

and it suffices to take $N = L$. Now suppose that $L = M$, so that

$$E(M) = E(L) = \wedge\{E(N) : N \in \mathcal{F},\ M \subsetneq N\}.$$

By Theorem 1.7.11, the set $\{E(N) : N \in \mathcal{F},\ M \subsetneq N\}$ has $E(M)$ in its strong closure, and therefore meets the strong neighbourhood U of $E(M)$. Hence there is a subspace N satisfying

$$N \in \mathcal{F}, \qquad M \subsetneq N, \qquad ||E(N)Ky - E(M)Ky|| < \tfrac{1}{2}\delta \quad (y \in S).$$

An argument, similar to the one used above to prove part (i) of the lemma, now shows that

$$||[E(N)-E(M)]K|| \leqslant \delta,$$

and this implies the inequality required in (ii).

Suppose that $\mathcal{F}$ is a simple chain of closed subspaces of a Hilbert space $\mathcal{H}$. By a *partition* of $\mathcal{F}$ we mean a finite subset $\mathcal{P}$ of $\mathcal{F}$ which contains the two trivial subspaces $\{0\}$ and $\mathcal{H}$. We may suppose that $\mathcal{P} = \{M_0, M_1, \ldots, M_n\}$, where

$$\{0\} = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_n = \mathcal{H},$$
$$M_j \in \mathcal{F} \quad (j = 0, 1, \ldots, n).$$

If $K \in \mathcal{B}(\mathcal{H})$, the *diagonal part of K relative to $\mathcal{P}$* is a bounded linear operator, denoted by $D(K, \mathcal{P})$ and defined by

(11) $$D(K, \mathcal{P}) = \sum_{j=1}^{n} [E(M_j)-E(M_{j-1})]K[E(M_j)-E(M_{j-1})].$$

If we write E_j for $E(M_j)-E(M_{j-1})$, then $E_1, \ldots, E_n$ are pairwise orthogonal projections with sum I. For each x in $\mathcal{H}$

$$\begin{aligned} ||D(K,\mathcal{P})x||^2 &= ||\sum_{j=1}^{n} E_jKE_jx||^2 \\ &= \sum_{j=1}^{n} ||E_jKE_jx||^2 \\ &\leqslant \sum_{j=1}^{n} ||E_jKE_j||^2||E_jx||^2 \\ &\leqslant [\max_{1\leqslant j\leqslant n} ||E_jKE_j||]^2 \sum_{j=1}^{n} ||E_jx||^2 \\ &= [\max_{1\leqslant j\leqslant n} ||E_jKE_j||]^2\ ||x||^2. \end{aligned}$$

Thus

(12) $$||D(K, \mathcal{P})|| \leqslant \max_{1\leqslant j\leqslant n} ||E_jKE_j|| \leqslant ||K||.$$

Now suppose that $\mathcal{P}_1$ is another partition of $\mathcal{F}$, consisting of the subspaces $\{0\} = N_0 \subsetneq N_1 \subsetneq \cdots \subsetneq N_m = \mathcal{H}$, and write F_k for $E(N_k)-E(N_{k-1})$ $(k = 1, \ldots, m)$. If $\mathcal{P} \subseteq \mathcal{P}_1$ then, for each $k = 1, 2, \ldots, m$, there is an integer r such that $1 \leqslant r \leqslant n$ and $M_{r-1} \subseteq N_{k-1} \subseteq N_k \subseteq M_r$. Thus $F_k \leqslant E_r$, and so $F_k E_j = E_j F_k = 0$ when $1 \leqslant j \leqslant n$ and $j \neq r$. It follows that

$$\begin{aligned} F_k K F_k &= F_k E_r K E_r F_k \\ &= F_k\left(\sum_{j=1}^{n} E_j K E_j\right)F_k = F_k D(K, \mathcal{P})F_k, \end{aligned}$$

and summation over $k = 1, \ldots, m$ yields

$$D(K, \mathcal{P}_1) = D(D(K, \mathcal{P}), \mathcal{P}_1) \quad \text{if } \mathcal{P} \subseteq \mathcal{P}_1.$$

This, together with (12), implies that

$$\|D(K, \mathcal{P}_1)\| \leqslant \|D(K, \mathcal{P})\| \quad \text{if } \mathcal{P} \subseteq \mathcal{P}_1. \tag{13}$$

LEMMA 4.4.4. *Suppose that $\mathcal{F}$ is a simple chain of closed subspaces of a Hilbert space $\mathcal{H}$, K is a compact linear operator acting on $\mathcal{H}$, $P(M)KP(M) = 0$ whenever $M \in \mathcal{F}$ and $M \neq M_-$, and δ is a positive real number. Then there is a partition $\mathcal{P}$ of $\mathcal{F}$ such that $\|D(K, \mathcal{P})\| \leqslant \delta$.*

Proof. We shall say that a subspace M in $\mathcal{F}$ is *satisfactory* if there exist $M_0, M_1, \ldots, M_r$ in $\mathcal{F}$ such that

$$\{0\} = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_r = M, \tag{14}$$

$$\|[E(M_j)-E(M_{j-1})]K[E(M_j)-E(M_{j-1})]\| \leqslant \delta \quad (j = 1, \ldots, r). \tag{15}$$

We have to show that $\mathcal{H}$ is satisfactory; for then, the subspaces $M_0, \ldots, M_r$ satisfying (14) and (15) form a partition $\mathcal{P}$ of $\mathcal{F}$, and $\|D(K, \mathcal{P})\| \leqslant \delta$ by (12). The proof that $\mathcal{H}$ is satisfactory is divided into three stages.

(a) *There exists N in $\mathcal{F}$ such that $N \neq \{0\}$ and N is satisfactory.* By Lemma 4.4.3 (ii), with $M = \{0\}$, there exists a non-zero subspace N in $\mathcal{F}$ such that $\|E(N)KE(N)\| \leqslant \delta$. By taking $M_0 = \{0\}$ and $M_1 = N$, it follows that N is satisfactory.

We now define a subspace M in $\mathcal{F}$ by

$$M = \vee\{N : N \in \mathcal{F},\ N \text{ is satisfactory}\}. \tag{16}$$

(b) *M is satisfactory.* It follows from part (a) of the proof that $M \neq \{0\}$, so Lemma 4.4.3 (i) implies the existence of a subspace L such that

$$L \in \mathcal{F}, \qquad L \subsetneq M, \quad \|[E(M)-E(L)]K[E(M)-E(L)]\| \leqslant \delta. \tag{17}$$

Since L is a proper subspace of M, we deduce from (16) that there is a satisfactory subspace N in $\mathcal{F}$ such that $L \subseteq N \subseteq M$. Hence there exist subspaces $M_0, \ldots, M_r$ in $\mathcal{F}$ such that

$$\{0\} = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_r = N,$$

and (15) is satisfied. Since $L \subseteq N \subseteq M$, $E(M)-E(N)$ is a subprojection of $E(M)-E(L)$ and so, by (17),

$$\|[E(M)-E(N)]K[E(M)-E(N)]\| \leqslant \delta.$$

By considering the subspaces $M_0, M_1, \ldots, M_r, M_{r+1}$, where $M_{r+1} = M$, it now follows that M is satisfactory.

(c) *$\mathcal{H}$ is satisfactory.* Suppose the contrary so that, by part (b) of the proof, $M \neq \mathcal{H}$. Since M is satisfactory, there exist subspaces $M_0, M_1, \ldots, M_r$ in $\mathcal{F}$ such that (14) and (15) are satisfied. By Lemma 4.4.3 (ii) there is a subspace M_{r+1} in $\mathcal{F}$ such that

$$M = M_r \subsetneq M_{r+1}, \quad \|[E(M_{r+1})-E(M_r)]K[E(M_{r+1})-E(M_r)]\| \leqslant \delta.$$

It is clear that M_{r+1} is satisfactory, and this contradicts (16) since $M \subsetneqq M_{r+1}$. Sence $\mathcal{H}$ is satisfactory, and the lemma is proved.

Suppose that $\mathcal{F}$ is a simple chain of closed subspaces of a Hilbert space $\mathcal{H}$, and $\mathcal{P}$ is a partition of $\mathcal{F}$, where

$$(18)\qquad \begin{cases} \mathcal{P} = \{M_0, M_1, \ldots, M_n\}, \\ \{0\} = M_0 \subsetneqq M_1 \subsetneqq \cdots \subsetneqq M_n = \mathcal{H}. \end{cases}$$

If $K \in \mathcal{B}(\mathcal{H})$, the bounded linear operator $S(K, \mathcal{P})$ defined by

$$(19)\qquad S(K,\mathcal{P}) = \sum_{j=1}^{n} E(M_{j-1})K[E(M_j)-E(M_{j-1})]$$

is called the *superdiagonal part of* K *relative to* $\mathcal{P}$. Clearly $S(K,\mathcal{P})(M_j) \subset M_{j-1}$ $(j = 1, \ldots, n)$. Given any non-zero subspace M in $\mathcal{F}$, there is an integer j such that $M_{j-1} \subsetneqq M \subseteq M_j$; thus $M_{j-1} \subseteq M_- \subseteq M \subseteq M_j$, and

$$S(K,\mathcal{P})(M) \subseteq S(K,\mathcal{P})(M_j) \subseteq M_{j-1} \subseteq M_-.$$

It follows that

$$(20)\qquad S(K,\mathcal{P})(M) \subseteq M_- \quad (M \in \mathcal{F}).$$

If we write E_j for $E(M_j)-E(M_{j-1})$, then the adjoint $S(K,\mathcal{P})^*$ is given by

$$\begin{aligned} S(K,\mathcal{P})^* &= \sum_{j=1}^{n} [E(M_j)-E(M_{j-1})]K^*E(M_{j-1}) \\ &= \sum_{j=2}^{n} E_jK^*\Big(\sum_{k=1}^{j-1} E_k\Big) \\ &= \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} E_jK^*E_k \\ &= \sum_{k=1}^{n-1} [I-E(M_k)]K^*[E(M_k)-E(M_{k-1})] \end{aligned}$$

$$= \sum_{j=1}^{n} [I-E(M_j)]K^*[E(M_j)-E(M_{j-1})].$$

If K is self-adjoint then

$$\begin{aligned} &S(K,\mathcal{P})+S(K,\mathcal{P})^* \\ &= \sum_{j=1}^{n} [I-E(M_j)+E(M_{j-1})]K[E(M_j)-E(M_{j-1})] \\ &= K-\sum_{j=1}^{n} [E(M_j)-E(M_{j-1})]K[E(M_j)-E(M_{j-1})] \\ &= K-D(K,\mathcal{P}), \end{aligned}$$

where $D(K,\mathcal{P})$ is the diagonal part of K relative to $\mathcal{P}$. Thus

$$(21)\qquad K = S(K,\mathcal{P})+S(K,\mathcal{P})^*+D(K,\mathcal{P})$$

if $K = K^* \in \mathcal{B}(\mathcal{H})$.

If $\mathcal{Q}$ is a collection of subspaces satisfying

$$(22)\qquad \begin{cases} \mathcal{Q} = \{L_1, \ldots, L_n\}, \qquad L_j \in \mathcal{F}, \\ M_{j-1} \subseteq L_j \subseteq M_j \quad (j = 1, \ldots, n), \end{cases}$$

then the ‘Riemann–Stieltjes sum’

$$(23)\qquad S(K,\mathcal{P},\mathcal{Q}) = \sum_{j=1}^{n} E(L_j)K[E(M_j)-E(M_{j-1})]$$

is a bounded linear operator on $\mathcal{H}$; the superdiagonal part $S(K,\mathcal{P})$ is an example of such a sum (obtained by taking $L_j = M_{j-1}$). We say that the integral

$$\int_{\mathcal{F}} E(M)KE(dM)$$

converges if there is a bounded linear operator A on $\mathcal{H}$ with the following property: given any positive real number δ, there is a partition $\mathcal{P}(\delta)$ of $\mathcal{F}$ such that

$$||A - S(K, \mathcal{P}, \mathcal{Q})|| \leqslant \delta$$

whenever the partition $\mathcal{P}$ of $\mathcal{F}$ contains $\mathcal{P}(\delta)$ (and, of course $\mathcal{P}$ and $\mathcal{Q}$ are related as in (18) and (22)). When this condition is satisfied, we write

$$A = \int_{\mathcal{F}} E(M)KE(dM), \tag{24}$$

and call A the *superdiagonal integral of K relative to $\mathcal{F}$*. By taking $\mathcal{Q} = \{M_0, \ldots, M_{n-1}'\}$, it follows that $||A - S(K, \mathcal{P})|| \leqslant \delta$ whenever $\mathcal{P}$ contains $\mathcal{P}(\delta)$; so A is a norm limit of operators of the form $S(K, \mathcal{P})$ and, by (20),

$$A(M) \subseteq M_- \quad (M \in \mathcal{F}). \tag{25}$$

LEMMA 4.4.5. *If $\mathcal{F}$ is a simple chain of closed subspaces of a Hilbert space $\mathcal{H}$, $K \in \mathcal{B}(\mathcal{H})$ and the superdiagonal integral $\int_{\mathcal{F}} E(M)KE(dM)$ converges, then $P(M)KP(M) = 0$ ($M \in \mathcal{F}$, $M \neq M_-$).*

Proof. Given any positive real number δ, let $\mathcal{P}(\delta)$ be a partition of $\mathcal{F}$ such that $||A - S(K, \mathcal{P}, \mathcal{Q})|| \leqslant \delta$ whenever $\mathcal{P}$ contains $\mathcal{P}(\delta)$, where A is the superdiagonal integral. If $M \in \mathcal{F}$ and $M \neq M_-$, let $\mathcal{P} = \{M_0, M_1, \ldots, M_n\}$ be the partition consisting of the members of $\mathcal{P}(\delta)$, together with M and M_-. Since M_- is the immediate predecessor of M in $\mathcal{F}$, there is an integer k such that $M_{k-1} = M_-$, $M_k = M$. We now consider the Riemann-Stieltjes sums $S(K, \mathcal{P}, \mathcal{Q}_1)$ and $S(K, \mathcal{P}, \mathcal{Q}_2)$, where $\mathcal{Q}_1 = \{M_1, \ldots, M_n\}$ and $\mathcal{Q}_2$ is obtained from $\mathcal{Q}_1$ by altering the kth. element to M_{k-1}. Since $||A - S(\mathcal{P}, K, \mathcal{Q}_j)|| \leqslant \delta$ ($j = 1, 2$) and

$$\begin{aligned} S(K, \mathcal{P}, \mathcal{Q}_1) - S(K, \mathcal{P}, \mathcal{Q}_2) &= [E(M_k) - E(M_{k-1})]K[E(M_k) - E(M_{k-1})] \\ &= P(M)KP(M), \end{aligned}$$

it follows that $||P(M)KP(M)|| \leqslant 2\delta$. This has been proved for every positive δ, so $P(M)KP(M) = 0$.

We can now prove the main result of this section.

THEOREM 4.4.6. *Suppose that $\mathcal{F}$ is a simple chain of closed subspaces of a Hilbert space $\mathcal{H}$, and K is a compact self-adjoint operator acting on $\mathcal{H}$. In order that there exist a quasi-nilpotent operator T which has skew-adjoint part K and leaves invariant each subspace in $\mathcal{F}$, it is necessary and sufficient that the superdiagonal integral of K relative to $\mathcal{F}$ should converge. When this is so, the unique operator T with the stated properties is given by*

$$T = 2i \int_{\mathcal{F}} E(M)KE(dM).$$

Proof. First, suppose that the superdiagonal integral converges, and denote its value by A. Then A is the limit, in norm, of operators of the form $S(K, \mathcal{P})$ defined by (19). Since each $S(K, \mathcal{P})$ is compact, so is A; and, by (25), $A(M) \subseteq M_-$ whenever $M \in \mathcal{F}$. It now follows from Theorem 4.3.10 that A is quasi-nilpotent.

Since the superdiagonal integral converges, Lemma 4.4.5 implies that $P(M)KP(M) = 0$ whenever $M \in \mathcal{F}$ and $M \neq M_-$. If δ is a positive real number then, by Lemma 4.4.4, there is a partition $\mathcal{P}_0$ of $\mathcal{F}$ such that $||D(K, \mathcal{P}_0)|| \leqslant \delta$; it follows from (13) that

$$||D(K, \mathcal{P})|| \leqslant \delta \quad \text{whenever } \mathcal{P}_0 \subseteq \mathcal{P}. \tag{26}$$

By definition of the superdiagonal integral, there is a partition $\mathcal{P}_1$ of $\mathcal{F}$ such that

$$||A - S(K, \mathcal{P})|| \leqslant \delta \quad \text{whenever } \mathcal{P}_1 \subseteq \mathcal{P}. \tag{27}$$

If $\mathcal{P}$ is a partition which contains both $\mathcal{P}_0$ and $\mathcal{P}_1$ then, by (21), (26) and (27),

$$\begin{aligned}||K-A-A^*|| &= ||[S(K,\mathcal{P})-A]+[S(K,\mathcal{P})-A]^*+D(K,\mathcal{P})||\\ &\leqslant 3\delta.\end{aligned}$$

Since this holds for every positive δ, it follows that $K = A+A^*$.

Since A is quasi-nilpotent, $A(M) \subseteq M_- \subseteq M$ whenever $M \in \mathcal{F}$, and $A+A^* = K$, the operator

$$T = 2iA = 2i\int_{\mathcal{F}} E(M)KE(dM)$$

is quasi-nilpotent, leaves each subspace in $\mathcal{F}$ invariant, and has skew-adjoint part K.

Conversely, suppose there is a quasi-nilpotent operator T such that $\operatorname{Im} T = K$ and $T(M) \subseteq M$ for each M in $\mathcal{F}$. By Theorem 4.4.2, T is compact, and $P(M)KP(M) = 0$ whenever $M \in \mathcal{F}$ and $M \neq M_-$. We may set $T = 2iA$, where A is compact and quasi-nilpotent, $A(M) \subseteq M$ for each M in $\mathcal{F}$, and $A+A^* = K$. By Lemma 4.4.1, $P(M)AP(M) = 0$ whenever $M \in \mathcal{F}$ and $M \neq M_-$. It now follows from Lemma 4.4.4 that, if δ is a positive real number, there exist partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ of $\mathcal{F}$ such that $||D(K,\mathcal{P}_1)|| \leqslant \delta$, $||D(A,\mathcal{P}_2)|| \leqslant \delta$. If $\mathcal{P}_\delta = \mathcal{P}_1 \cup \mathcal{P}_2$ then, by (13),

$$(28) \qquad ||D(K,\mathcal{P})|| \leqslant \delta, \quad ||D(A,\mathcal{P})|| \leqslant \delta$$

whenever $\mathcal{P}_\delta \subseteq \mathcal{P}$.

Suppose that $\mathcal{P} = \{M_0, M_1, \ldots, M_n\}$ is a partition of $\mathcal{F}$, and $\mathcal{Q} = \{L_1, \ldots, L_n\}$ satisfies the usual condition (22). If we write E_j and F_j for the projections $E(M_j)-E(M_{j-1})$ and $E(L_j)-E(M_{j-1})$ respectively, then $0 \leqslant F_j \leqslant E_j$, while $E_1, \ldots, E_n$ are pairwise orthogonal; so the projection $F = F_1+\ldots+F_n$ satisfies $FE_j = F_j$ $(j = 1, \ldots, n)$. It follows from (19) and (23) that

$$\begin{aligned}S(K,\mathcal{P},\mathcal{Q})-S(K,\mathcal{P}) &= \sum_{j=1}^{n} [E(L_j)-E(M_{j-1})]K[E(M_j)-E(M_{j-1})]\\ &= \sum_{j=1}^{n} F_jKE_j\\ &= F\sum_{j=1}^{n} E_jKE_j = FD(K,\mathcal{P}).\end{aligned}$$

Thus

$$(29) \qquad ||S(K,\mathcal{P},\mathcal{Q})-S(K,\mathcal{P})|| \leqslant ||D(K,\mathcal{P})||.$$

In order to estimate $||A-S(K,\mathcal{P})||$, we first note that

$$(30) \qquad E(M_{j-1})A^*[E(M_j)-E(M_{j-1})] = 0 \quad (j = 1, \ldots, n),$$

$$(31) \qquad [I-E(M_j)]A[E(M_j)-E(M_{j-1})] = 0 \quad (j = 1, \ldots, n).$$

Indeed, since the subspaces M_j and M_{j-1} are invariant under A, the left-hand side of (30) is the adjoint of

$$[E(M_j)-E(M_{j-1})]E(M_{j-1})AE(M_{j-1}) = 0,$$

while the left-hand side of (31) is

$$\begin{aligned}&[I-E(M_j)]AE(M_j)[E(M_j)-E(M_{j-1})]\\ &\quad = [I-E(M_j)]E(M_j)AE(M_j)[E(M_j)-E(M_{j-1})] = 0.\end{aligned}$$

Since $K = A+A^*$, it follows from (19), (30) and (31) that

$$\begin{aligned}S(K,\mathcal{P}) &= \sum_{j=1}^{n} E(M_{j-1})(A+A^*)[E(M_j)-E(M_{j-1})]\\ &= \sum_{j=1}^{n} E(M_{j-1})A[E(M_j)-E(M_{j-1})]\\ &= \sum_{j=1}^{n} [I-E(M_j)+E(M_{j-1})]A[E(M_j)-E(M_{j-1})]\end{aligned}$$

$$= A - \sum_{j=1}^{n} [E(M_j)-E(M_{j-1})]A[E(M_j)-E(M_{j-1})]$$
$$= A - D(A, \mathcal{P}).$$

Thus $A - S(K, \mathcal{P}) = D(A, \mathcal{P})$ and, by (29) and (28),

$$\|A - S(K, \mathcal{P}, \mathcal{Q})\| \leqslant \|A - S(K, \mathcal{P})\| + \|S(K, \mathcal{P}) - S(K, \mathcal{P}, \mathcal{Q})\|$$
$$\leqslant \|D(A, \mathcal{P})\| + \|D(K, \mathcal{P})\| \leqslant 2\delta,$$

whenever the partition $\mathcal{P}$ contains $\mathcal{P}_\delta$. This shows that the superdiagonal integral of K relative to $\mathcal{F}$ converges, and that its value is A. Since

$$T = 2iA = 2i \int_{\mathcal{F}} E(M)KE(dM),$$

the theorem is proved.

In view of Theorem 4.4.6, it is desirable to find conditions on K which ensure the convergence of the superdiagonal integral $\int E(M)KE(dM)$. Before proving our first result in this direction, Theorem 4.4.10, we require some lemmas. As in Chapter 2, $\mathcal{C}_p$ denotes the von Neumann–Schatten class of operators, $\|\ \|_p$ is the usual norm on $\mathcal{C}_p$, and τ is the trace on $\mathcal{C}_1$. We recall from §2.4 that the Schmidt class $\mathcal{C}_2$ is a Hilbert space, with inner product $[\ ,\ ]$ defined by $[S, T] = \tau(T^*S)$.

LEMMA 4.4.7. *Suppose that $\mathcal{F}$ is a simple chain of closed subspaces of a Hilbert space $\mathcal{H}$, K is a self-adjoint Schmidt class operator on $\mathcal{H}$, $\mathcal{P} = \{M_0, M_1, \ldots, M_n\}$ is a partition of $\mathcal{F}$ and $\mathcal{Q} = \{L_1, \ldots, L_n\}$ satisfies the usual condition (22). Then the operators $D(K, \mathcal{P})$, $S(K, \mathcal{P})$ and $S(K, \mathcal{P}, \mathcal{Q})$, defined by (11), (19) and (23), all lie in the Schmidt class, and*

(i) $\quad \|S(K, \mathcal{P}, \mathcal{Q}) - S(K, \mathcal{P})\|_2 \leqslant \|D(K, \mathcal{P})\|_2,$

(ii) $\quad \|D(K, \mathcal{P})\|_2^2 + \|K - D(K, \mathcal{P})\|_2^2 = \|K\|_2^2.$

Proof. Since $\mathcal{C}_2$ is an ideal in $\mathcal{B}(\mathcal{H})$ and $K \in \mathcal{C}_2$, it is apparent that $D(K, \mathcal{P})$, $S(K, \mathcal{P})$, $S(K, \mathcal{P}, \mathcal{Q}) \in \mathcal{C}_2$. We have seen, in proving Theorem 4.4.6, that

$$S(K, \mathcal{P}, \mathcal{Q}) - S(K, \mathcal{P}) = FD(K, \mathcal{P}),$$

where F is the projection $\Sigma\{E(L_j) - E(M_{j-1})\}$; thus

$$\|S(K, \mathcal{P}, \mathcal{Q}) - S(K, \mathcal{P})\|_2 \leqslant \|F\|_\infty \|D(K, \mathcal{P})\|_2 = \|D(K, \mathcal{P})\|_2.$$

If we write E_j for the projection $E(M_j) - E(M_{j-1})$, then $E_1, \ldots, E_n$ are pairwise orthogonal and have sum I. Note that

$$[E_rKE_s, E_pKE_q] = \tau((E_pKE_q)^*(E_rKE_s))$$
$$= \tau(E_qKE_pE_rKE_s)$$
$$= \tau(KE_pE_rKE_sE_q),$$

and the right-hand side is 0 unless $p = r$ and $q = s$. Hence the operators $\{E_rKE_s : r, s = 1, \ldots, n\}$ are pairwise orthogonal vectors in the Hilbert space $\mathcal{C}_2$. It follows that the vectors

$$D(K, \mathcal{P}) = \sum_{r=1}^{n} E_rKE_r,$$

$$K - D(K, P) = \sum_{r \neq s} E_rKE_s$$

in $\mathcal{C}_2$ are orthogonal, and therefore

$$\|K\|_2^2 = \|D(K, \mathcal{P}) + \{K - D(K, \mathcal{P})\}\|_2^2$$
$$= \|D(K, \mathcal{P})\|_2^2 + \|K - D(K, \mathcal{P})\|_2^2.$$

LEMMA 4.4.8. *Suppose that $\mathcal{F}$ is a simple chain of closed subspaces of a Hilbert space $\mathcal{H}$, K is a self-adjoint Schmidt class operator acting on $\mathcal{H}$, $P(M)KP(M) = 0$ whenever $M \in \mathcal{F}$ and $M \neq M_-$, and δ is a positive real number. Then there is a partition $\mathcal{P}_0$ of $\mathcal{F}$ with the following property: if $\mathcal{P}$ is a partition of $\mathcal{F}$ which contains $\mathcal{P}_0$, then $||D(K,\mathcal{P})||_2 < \delta$.*

Proof. If $\{\varphi_j : j \in J\}$ is an orthonormal basis of $\mathcal{H}$, then

$$||K||_2^2 = \sum_{j,k \in J} |\langle K\varphi_j, \varphi_k\rangle|^2,$$

so there is a finite subset F of J such that

$$||K||_2^2 < \sum_{j,k \in F} |\langle K\varphi_j, \varphi_k\rangle|^2 + \tfrac{1}{2}\delta^2.$$

Given any partition $\mathcal{P}$ of $\mathcal{F}$,

$$\begin{aligned} &||K-D(K,\mathcal{P})||_2^2 \\ &\geqslant \sum_{j,k \in F} |\langle K\varphi_j, \varphi_k\rangle - \langle D(K,\mathcal{P})\varphi_j, \varphi_k\rangle|^2 \\ &= \sum_{j,k \in F} \{|\langle K\varphi_j, \varphi_k\rangle|^2 + |\langle D(K,\mathcal{P})\varphi_j, \varphi_k\rangle|^2 \\ &\qquad -2\mathrm{Re}[\langle K\varphi_j, \varphi_k\rangle \langle \varphi_k, D(K,\mathcal{P})\varphi_j\rangle]\} \\ &\geqslant \sum_{j,k \in F} \{|\langle K\varphi_j, \varphi_k\rangle|^2 - 2||K||\ ||D(K,\mathcal{P})||\} \\ &> ||K||_2^2 - \tfrac{1}{2}\delta^2 - 2m^2||K||\ ||D(K,\mathcal{P})||, \end{aligned}$$

where m is the number of elements in F. By Lemma 4.4.7 (ii) the above inequality is equivalent to

$$(32) \qquad ||D(K,\mathcal{P})||_2^2 < \tfrac{1}{2}\delta^2 + 2m^2||K||\ ||D(K,\mathcal{P})||.$$

Let η be a positive real number such that $2m^2||K||\eta < \tfrac{1}{2}\delta^2$. By Lemma 4.4.4, there is a partition $\mathcal{P}_0$ of $\mathcal{F}$ such that $||D(K,\mathcal{P}_0)|| \leqslant \eta$.

If $\mathcal{P}$ is any partition of $\mathcal{F}$ which contains $\mathcal{P}_0$, then $||D(K,\mathcal{P})|| \leqslant \eta$ by (13); and it then follows from (32) that

$$||D(K,\mathcal{P})||_2^2 < \tfrac{1}{2}\delta^2 + 2m^2||K||\eta < \delta^2.$$

LEMMA 4.4.9. *Suppose that $\mathcal{F}$ is a simple chain of closed subspaces of a Hilbert space $\mathcal{H}$, K is a self-adjoint Schmidt class operator on $\mathcal{H}$, and $\mathcal{P}_1$, $\mathcal{P}_2$ are partitions of $\mathcal{F}$. Then*

$$2||S(K,\mathcal{P}_1)-S(K,\mathcal{P}_2)||_2^2 = ||D(K,\mathcal{P}_1)-D(K,\mathcal{P}_2)||_2^2.$$

Proof. By (21),

$$D(K,\mathcal{P}_r) = K - S(K,\mathcal{P}_r) - S(K,\mathcal{P}_r)^* \quad (r = 1, 2),$$

so

$$\begin{aligned} &D(K,\mathcal{P}_1)-D(K,\mathcal{P}_2) \\ &\qquad = \{S(K,\mathcal{P}_2)-S(K,\mathcal{P}_1)\} + \{S(K,\mathcal{P}_2)-S(K,\mathcal{P}_1)\}^*. \end{aligned}$$

If we show that the two operators on the right-hand side of the last equation are orthogonal vectors in the Hilbert space $\mathcal{C}_2$, it follows that

$$\begin{aligned} &||D(K,\mathcal{P}_1)-D(K,\mathcal{P}_2)||_2^2 \\ &\qquad = ||S(K,\mathcal{P}_2)-S(K,\mathcal{P}_1)||_2^2 + ||S(K,\mathcal{P}_2)^*-S(K,\mathcal{P}_1)^*||_2^2, \end{aligned}$$

and we obtain the required result, since $||T^*||_2 = ||T||_2$ when $T \in \mathcal{C}_2$. It is therefore sufficient to prove that

$$[S(K,\mathcal{P}_r),\ S(K,\mathcal{P}_s)^*] = 0 \quad (r, s = 1, 2);$$

that is, $\tau(S(K,\mathcal{P}_s)S(K,\mathcal{P}_r)) = 0 \quad (r, s = 1, 2)$.

Suppose that $\mathcal{P}_r$ consists of the subspaces

$$\{0\} = M_0 \subsetneqq M_1 \subsetneqq \dots \subsetneqq M_n = \mathcal{H}$$

(in $\mathcal{F}$). It follows at once from the definition (19) of $S(K, \mathcal{P}_r)$ that $S(K, \mathcal{P}_r)$ maps M_j into M_{j-1} ($j = 1, \ldots, n$); moreover, by (20), $S(K, \mathcal{P}_s)$ leaves invariant each subspace in $\mathcal{F}$. Hence the operator $T = S(K, \mathcal{P}_s)S(K, \mathcal{P}_r)$ satisfies

$$T(M_j) \subseteq M_{j-1} \quad (j = 1, \ldots, n). \tag{33}$$

The range of the projection $E_j = E(M_j)-E(M_{j-1})$ is contained in M_j, so (33) implies that $TE_j = E(M_{j-1})TE_j$. It follows that

$$T = \sum_{j=1}^{n} TE_j = \sum_{j=1}^{n} E(M_{j-1})TE_j;$$

and, since $E_jE(M_{j-1}) = 0$,

$$\begin{aligned} \tau(S(K, \mathcal{P}_s)S(K, \mathcal{P}_r)) &= \tau(T) \\ &= \sum_{j=1}^{n} \tau(E(M_{j-1})TE_j) \\ &= \sum_{j=1}^{n} \tau(E_jE(M_{j-1})T) = 0. \end{aligned}$$

THEOREM 4.4.10. *Suppose that $\mathcal{F}$ is a simple chain of closed subspaces of a Hilbert space $\mathcal{H}$, K is a self-adjoint Schmidt class operator acting on $\mathcal{H}$, and $P(M)KP(M) = 0$ whenever $M \in \mathcal{F}$ and $M \neq M_-$. Then the superdiagonal integral of K relative to $\mathcal{F}$ converges; the operator*

$$A = \int_{\mathcal{F}} E(M)KE(dM)$$

lies in the Schmidt class $\mathcal{C}_2$, and the approximating Riemann–Stieltjes sums $S(K, \mathcal{P}, \mathcal{Q})$ converge to A with respect to the norm $\|\ \|_2$ on $\mathcal{C}_2$.

Proof. Suppose that, for $r = 1, 2$, $\mathcal{P}_r$ is a partition of $\mathcal{F}$, and $S(K, \mathcal{P}_r, \mathcal{Q}_r)$ is a Riemann–Stieltjes sum of the type (23); of course, we assume that $\mathcal{P}_r$ and $\mathcal{Q}_r$ are related as in (18) and (22). By Lemmas 4.4.7(i) and 4.4.9,

$$\begin{aligned} &\|S(K, \mathcal{P}_1, \mathcal{Q}_1)-S(K, \mathcal{P}_2, \mathcal{Q}_2)\|_2 \\ &\quad \leqslant \|S(K, \mathcal{P}_1, \mathcal{Q}_1)-S(K, \mathcal{P}_1)\|_2+\|S(K, \mathcal{P}_1)-S(K, \mathcal{P}_2)\|_2 + \\ &\qquad +\|S(K, \mathcal{P}_2)-S(K, \mathcal{P}_2, \mathcal{Q}_2)\|_2 \\ &\quad \leqslant \|D(K, \mathcal{P}_1)\|_2 + 2^{-1/2}\|D(K, \mathcal{P}_1)-D(K, \mathcal{P}_2)\|_2+\|D(K, \mathcal{P}_2)\|_2. \end{aligned}$$

Thus

$$\|S(K, \mathcal{P}_1, \mathcal{Q}_1)-S(K, \mathcal{P}_2, \mathcal{Q}_2)\| \leqslant 2\|D(K, \mathcal{P}_1)\|+2\|D(K, \mathcal{P}_2)\|. \tag{34}$$

Given a positive real number δ, let $\mathcal{P}_0$ be a partition of $\mathcal{F}$ which satisfies the conclusions of Lemma 4.4.8. If $\mathcal{P}_1$ and $\mathcal{P}_2$ both contain $\mathcal{P}_0$, then $\|D(K, \mathcal{P}_r)\|_2 < \delta$ $(r = 1, 2)$; and (34) gives

$$\|S(K, \mathcal{P}_1, \mathcal{Q}_1)-S(K, \mathcal{P}_2, \mathcal{Q}_2)\| < 4\delta \text{ if } \mathcal{P}_0 \subseteq \mathcal{P}_1,\ \mathcal{P}_0 \subseteq \mathcal{P}_2. \tag{35}$$

The set of all pairs $(\mathcal{P}, \mathcal{Q})$ is directed by the inclusion relation (applied to $\mathcal{P}$) and the last inequality shows that the Riemann–Stieltjes sums $S(K, \mathcal{P}, \mathcal{Q})$ form a Cauchy net in the Hilbert space $\mathcal{C}_2$. Accordingly, this net converges to an element A of $\mathcal{C}_2$, with respect to $\|\ \|_2$ (hence, also, with respect to the norm on $\mathcal{B}(\mathcal{H})$).

In Theorem 4.5.7, we shall prove the convergence of the superdiagonal integral under conditions much less restrictive than those assumed in Theorem 4.4.10.

COROLLARY 4.4.11. *If T is a quasi-nilpotent operator acting on a Hilbert space $\mathcal{H}$, and* Im T *lies in the Schmidt class $\mathcal{C}_2$, then $T \in \mathcal{C}_2$.*

Proof. Since Im T is compact, it follows from Theorem 1.8.8 that T is compact. By Theorems 4.3.4, and 4.4.6, we deduce that

there is a simple chain $\mathcal{F}$ of closed subspaces of $\mathcal{H}$, each of which is invariant under T; and

$$T = 2i\int_{\mathcal{F}} E(M)KE(dM),$$

where K is the Schmidt class operator Im T. The convergence of the superdiagonal integral implies that $P(M)KP(M) = 0$ whenever $M \in \mathcal{F}$ and $M \neq M_-$, by Lemma 4.4.5, so the required result follows from Theorem 4.4.10.

4.5. *The operator of integration in* $L_2(0,1)$

This section is concerned with the operator V, acting on the Hilbert space $L_2(0,1)$ (Lebesgue measure) which is defined by the equation

(1)
$$(Vx)(s) = i\int_0^s x(t)dt$$
$$(x \in L_2(0,1) : 0 \leqslant s \leqslant 1).$$

Since V is a Volterra integral operator, with kernel h defined by

$$h(s, t) = \begin{cases} 0 & (0 \leqslant s < t \leqslant 1), \\ 1 & (0 \leqslant t \leqslant s \leqslant 1), \end{cases}$$

the discussion in §4.3.12 shows that V is quasi-nilpotent. The adjoint operator V^* arises from the kernel $\overline{h(t, s)}$, and is therefore given by

(2)
$$(V^*x)(s) = -i\int_s^1 x(t)dt.$$

Hence

$$(Vx)(s)-(V^*x)(s) = i\int_0^1 x(t)dt$$
$$= i\langle x, e\rangle e(s),$$

where e is the unit vector in $L_2(0,1)$ defined by

(3)
$$e(s) = 1 \quad (0 \leqslant s \leqslant 1).$$

Thus

(4)
$$(V-V^*)x = i\langle x, e\rangle e \quad (x \in L_2(0,1)).$$

It is readily verified that

$$(V^n e)(s) = i^n s^n/n! \quad (n = 0, 1, 2, \ldots : 0 \leqslant s \leqslant 1).$$

Since polynomials form an everywhere dense subspace of $L_2(0,1)$,

(5)
$$\mathfrak{L}(e, Ve, V^2e, \ldots) = L_2(0,1)$$

(here, as in §4.2, $\mathfrak{L}(x_1, x_2, \ldots)$ denotes the closed subspace generated by $x_1, x_2, \ldots$). Finally, we recall from §4.3.12 that, if

(6)
$$L_\lambda = \{x \in L_2(0,1) : x(s) = 0 \text{ almost everywhere on } [0,\lambda]\}$$

when $0 \leqslant \lambda \leqslant 1$, then $\mathcal{F} = \{L_\lambda : 0 \leqslant \lambda \leqslant 1\}$ is a continuous chain of closed subspaces of $L_2(0,1)$, each of which is invariant under V.

The preceding discussion shows that V is an example of a linear operator T, acting on a Hilbert space $\mathcal{H}$, which has the following properties:

(a) T is quasi-nilpotent,

(b) there is a unit vector e in $\mathcal{H}$ such that

$$(T-T^*)x = i\langle x, e\rangle e \quad (x \in \mathcal{H}),$$

(c) $\mathfrak{L}(e, Te, T^2e, \ldots) = \mathcal{H}$.

We shall prove, in Theorems 4.5.3 and 4.5.4, that every such operator T is unitarily equivalent to V, and that the subspaces L_λ $(0 \leqslant \lambda \leqslant 1)$ are the only closed invariant subspaces of V. These results are then used to obtain a condition, much more general than that of Theorem 4.4.10, which ensures the convergence of a superdiagonal integral.

Suppose that T is a bounded linear operator acting on a Hilbert space $\mathcal{H}$, and conditions (a), (b) and (c) above are satisfied. In view of (b), the skew-adjoint part K of T is given by

$$Kx = \tfrac{1}{2}\langle x, e\rangle e \quad (x \in \mathcal{H}). \tag{7}$$

Since K is compact, it follows from Theorem 1.8.8 that T is compact, so there is a simple chain $\mathcal{F}$ of closed subspaces of $\mathcal{H}$, each of which is invariant under T. By Theorem 4.4.6

$$T = 2i\int_{\mathcal{F}} E(M)KE(dM). \tag{8}$$

LEMMA 4.5.1. *Suppose that $\mathcal{H}$ is a Hilbert space, $T \in \mathcal{B}(\mathcal{H})$, and T satisfies conditions (a), (b) and (c). Let $\mathcal{F}$ be a simple chain of closed subspaces of $\mathcal{H}$, each of which is invariant under T. Then*

(i) the closed subspace generated by the vectors $E(M)e$ $(M \in \mathcal{F})$ is $\mathcal{H}$;

(ii) $\mathcal{F}$ is a continuous chain;

(iii) if $\varphi(M) = \langle E(M)e, e\rangle$ $(M \in \mathcal{F})$, then φ is a one-to-one mapping from $\mathcal{F}$ onto $[0,1]$, and $\varphi(M_1) \leqslant \varphi(M_2)$ whenever $M_1, M_2 \in \mathcal{F}$ and $M_1 \subseteq M_2$.

Proof. (i) Let X be the closed subspace generated by the vectors $E(M)e$ $(M \in \mathcal{F})$. If $x \in \mathcal{H}$ and $S(K, \mathcal{P}, \mathcal{Q})$ is a Riemann-Stieltjes sum of the type defined in 4.4(23), it follows from (7) that

$$\begin{aligned} S(K, \mathcal{P}, \mathcal{Q})x &= \sum_{j=1}^{n} E(L_j)K[E(M_j)-E(M_{j-1})]x \\ &= \sum_{j=1}^{n} \tfrac{1}{2}\langle [E(M_j)-E(M_{j-1})]x, e\rangle E(L_j)e \\ &\in X. \end{aligned}$$

By (8), T is the limit, in norm, of operators $S(K, \mathcal{P}, \mathcal{Q})$; so Tx is a norm limit of vectors $S(K, \mathcal{P}, \mathcal{Q})x$ in X, and therefore $Tx \in X$.

Since $e = E(\mathcal{H})e \in X$, and X contains $Te, T^2e, \ldots$ by the argument of the preceding paragraph, it follows from property (c) that $X = \mathcal{H}$.

(ii), (iii) For each M in $\mathcal{F}$, $\varphi(M) = \langle E(M)e, e\rangle = \|E(M)e\|^2$, so $0 \leqslant \varphi(M) \leqslant \|e\|^2 = 1$. Suppose that $M_1, M_2 \in \mathcal{F}$ and $M_1 \subseteq M_2$. Then

$$\varphi(M_2)-\varphi(M_1) = \langle [E(M_2)-E(M_1)]e, e\rangle = \|[E(M_2)-E(M_1)]e\|^2 \geqslant 0.$$

If $\varphi(M_1) = \varphi(M_2)$ then $E(M_1)e = E(M_2)e$ and, for each M in $\mathcal{F}$,

$$E(M_1)E(M)e = E(M)E(M_1)e = E(M)E(M_2)e = E(M_2)E(M)e.$$

It now follows from part (i) of the lemma that $E(M_1) = E(M_2)$, so $M_1 = M_2$. This proves that φ is a one-to-one order preserving mapping from $\mathcal{F}$ into $[0,1]$.

Suppose that $M \in \mathcal{F}$. If $M \neq M_-$ then $P(M)KP(M) = 0$ by Theorem 4.4.2, and it follows from (7) that

$$0 = P(M)KP(M)e = \tfrac{1}{2}\langle P(M)e, e\rangle P(M)e = \tfrac{1}{2}\|P(M)e\|^2 P(M)e.$$

Thus $P(M)e = 0$, and

$$\varphi(M)-\varphi(M_-) = \langle [E(M)-E(M_-)]e, e\rangle = \langle P(M)e, e\rangle = 0,$$

contradicting our earlier conclusion that φ is one-to-one. Thus $M = M_-$ for each M in $\mathcal{F}$; that is, $\mathcal{F}$ is a continuous chain.

It remains to prove that φ maps $\mathcal{F}$ *onto* $[0,1]$. Now $\varphi(\{0\}) = 0$ and $\varphi(\mathcal{H}) = 1$. If $0 < t < 1$, let

$$M = \cap\{L : L \in \mathcal{F}, \|E(L)e\|^2 \geqslant t\}. \tag{9}$$

Then

$$E(M) = \wedge\{E(L) : L \in \mathcal{F}, \|E(L)e\|^2 \geqslant t\},$$

and $M \in \mathcal{F}$ since $\mathcal{F}$ is a simple chain. By Theorem 1.7.11, $E(M)$ lies in the strong closure of $\{E(L) : L \in \mathcal{F}, \|E(L)e\|^2 \geqslant t\}$; since the

mapping $A \to ||Ae||^2$ is continuous for the strong topology on $\mathcal{B}(\mathcal{H})$, it follows that $||E(M)e||^2 \geqslant t$. Hence $M \neq \{0\}$ and, since $\mathcal{F}$ is continuous,

$$E(M) = E(M_-) = \vee\{E(L) : L \in \mathcal{F},\ L \subsetneq M\}.$$

Thus $E(M)$ lies in the strong closure of $\{E(L) : L \in \mathcal{F}, L \subsetneq M\}$. It follows from (9) that $||E(L)e||^2 < t$ whenever $L \in \mathcal{F}$ and $L \subsetneq M$; so the strong continuity of the mapping $A \to ||Ae||^2$ now implies that $||E(M)e||^2 \leqslant t$. Since the reverse inequality has already been established, $||E(M)e||^2 = t$; that is, $\varphi(M) = t$.

In proving our next lemma, we shall make use of some simple properties of an isometric isomorphism from a Hilbert space $\mathcal{H}_1$ onto another such space $\mathcal{H}_2$. Since $\langle x, x\rangle = \langle Ux, Ux\rangle$ for each x in $\mathcal{H}_1$, it follows immediately from the identity 1.7(2) that $\langle x, y\rangle = \langle Ux, Uy\rangle$ whenever $x,y \in \mathcal{H}_1$. Furthermore, if $T \in \mathcal{B}(\mathcal{H})$ and $x,y \in \mathcal{H}_1$, then

$$\begin{aligned}\langle (UTU^{-1})^*Ux,\ Uy\rangle &= \langle Ux,\ UTU^{-1}Uy\rangle\\ &= \langle Ux,\ UTy\rangle\\ &= \langle x,\ Ty\rangle\\ &= \langle T^*x,\ y\rangle\\ &= \langle UT^*x,\ Uy\rangle\\ &= \langle (UT^*U^{-1})Ux,\ Uy\rangle;\end{aligned}$$

so $(UTU^{-1})^* = UT^*U^{-1}$.

LEMMA 4.5.2. *Suppose that, for* $r = 1,2$, $\mathcal{H}_r$ *is a Hilbert space,* $T_r \in \mathcal{B}(\mathcal{H})$, *and conditions (a), (b) and (c) are satisfied with* T_r, $\mathcal{H}_r$, e_r *in place of* T, $\mathcal{H}$, e. *Let* $\mathcal{F}_r$ *be a simple chain of closed subspaces of* $\mathcal{H}_r$, *each of which is invariant under* T_r. *Then there is an isometric isomorphism* U *from* $\mathcal{H}_1$ *onto* $\mathcal{H}_2$ *such that*

$$Ue_1 = e_2, \qquad UT_1U^{-1} = T_2,$$
$$\{U(M) : M \in \mathcal{F}_1\} = \mathcal{F}_2.$$

Proof. By Lemma 4.5.1 (iii), the equation $\varphi_r(M) = \langle E(M)e_r,\ e_r\rangle$ defines a one-to-one mapping φ_r from $\mathcal{F}_r$ onto $[0,1]$, and $\varphi_r(M) \leqslant \varphi_r(N)$ if $M, N \in \mathcal{F}_r$ and $M \subseteq N$. The composite mapping $\varphi = \varphi_2^{-1} \circ \varphi_1$ is one-to-one, from $\mathcal{F}_1$ onto $\mathcal{F}_2$, and $\varphi(M) \subseteq \varphi(N)$ whenever $M, N \in \mathcal{F}_1$ and $M \subseteq N$; thus $\varphi(\{0\}) = \{0\}$, $\varphi(\mathcal{H}_1) = \mathcal{H}_2$ and

$$\varphi(M \wedge N) = \varphi(M) \wedge \varphi(N) \quad (M, N \in \mathcal{F}_1).$$

Furthermore

$$\begin{aligned}\langle E(\varphi(M))e_2,\ e_2\rangle &= \varphi_2(\varphi(M)) = \varphi_1(M)\\ &= \langle E(M)e_1,\ e_1\rangle \quad (M \in \mathcal{F}_1).\end{aligned}$$

It follows that, if $M, N \in \mathcal{F}_1$, then

$$\begin{aligned}\langle E(\varphi(M))e_2,\ E(\varphi(N))e_2\rangle &= \langle E(\varphi(N))E(\varphi(M))e_2,\ e_2\rangle\\ &= \langle E(\varphi(N) \wedge \varphi(M))e_2,\ e_2\rangle\\ &= \langle E(\varphi(N \wedge M))e_2,\ e_2\rangle\\ &= \langle E(N \wedge M)e_1,\ e_1\rangle\\ &= \langle E(N)E(M)e_1,\ e_1\rangle\\ &= \langle E(M)e_1,\ E(N)e_1\rangle.\end{aligned}$$

By linearity of the inner products, if α_j, β_k are scalars and M_j, $N_k \in \mathcal{F}_1$ $(1 \leqslant j \leqslant m,\ 1 \leqslant k \leqslant n)$, then

$$\begin{aligned}&\langle \Sigma\, \alpha_j E(\varphi(M_j))e_2,\ \Sigma\, \beta_k E(\varphi(N_k))e_2\rangle\\ &\qquad = \langle \Sigma\, \alpha_j E(M_j)e_1,\ \Sigma\, \beta_k E(N_k)e_2\rangle.\end{aligned}$$

In particular

$$\|\Sigma\, \alpha_j E(\varphi(M_j))e_2\| = \|\Sigma\, \alpha_j E(M_j)e_1\|. \tag{10}$$

For $r = 1, 2$, let X_r denote the set of all finite linear combinations of vectors of the form $E(M)e_r$, with M in $\mathcal{F}_r$; by Lemma 4.5.1 (i), X_r is an everywhere dense subspace of $\mathcal{H}_r$. If $x \in X_1$, and x is represented in two ways,

$$x = \sum_{j=1}^{m} \alpha_j E(M_j)e_1 = \sum_{k=1}^{n} \beta_k E(N_k)e_1,$$

as a linear combination of vectors $E(M)e_1$ $(M \in \mathcal{F}_1)$, then by (10),

$$\begin{aligned}\|\sum_{j=1}^{m} \alpha_j E(\varphi(M_j))e_2 - \sum_{k=1}^{n} \beta_k E(\varphi(N_k))e_2\| \\ = \|\sum_{j=1}^{m} \alpha_j E(M_j)e_1 - \sum_{k=1}^{n} \beta_k E(N_k)e_1\| = \|x-x\| = 0;\end{aligned}$$

thus

$$\sum_{j=1}^{m} \alpha_j E(\varphi(M_j))e_2 = \sum_{k=1}^{n} \beta_k E(\varphi(N_k))e_2.$$

It follows that the equation

$$U(\sum_{j=1}^{m} \alpha_j E(M_j)e_1) = \sum_{j=1}^{m} \alpha_j E(\varphi(M_j))e_2$$

defines (unambiguously) an isometric isomorphism U from X_1 onto X_2; this extends by continuity to an isometric isomorphism, which we denote by the same symbol U, from $\mathcal{H}_1$ onto $\mathcal{H}_2$.

Since $\varphi(\mathcal{H}_1) = \mathcal{H}_2$ and $e_r = E(\mathcal{H}_r)e_r$, it follows that $Ue_1 = e_2$. When $M, N \in \mathcal{F}_1$,

$$\begin{aligned} UE(M)E(N)e_1 &= UE(M \wedge N)e_1 \\ &= E(\varphi(M \wedge N))e_2 \\ &= E(\varphi(M) \wedge \varphi(M))e_2 \\ &= E(\varphi(M))E(\varphi(N))e_2 \\ &= E(\varphi(M))UE(N)e_1. \end{aligned}$$

By linearity, $UE(M)x = E(\varphi(M))Ux$ whenever $x \in X_1$; since X_1 is dense in $\mathcal{H}_1$, it follows that $UE(M) = E(\varphi(M))U$. Thus $E(\varphi(M)) = UE(M)U^{-1}$, and $\varphi(M) = U(M)$, for each M in $\mathcal{F}_1$; so

$$\mathcal{F}_2 = \{\varphi(M) : M \in \mathcal{F}_1\} = \{U(M) : M \in \mathcal{F}_1\}.$$

If $T_3 = UT_1U^{-1}$, then T is a quasi-nilpotent operator on $\mathcal{H}_2$, and we have to show that $T_3 = T_2$. If $N \in \mathcal{F}_2$ then $N = U(M)$ for some M in $\mathcal{F}_1$; thus

$$T_3(N) = UT_1U^{-1}U(M) = UT_1(M) \subseteq U(M) = N.$$

Since T_1 and T_2 both have property (b),

$$\begin{aligned}(T_3 - T_3^*)x &= U(T_1 - T_1^*)U^{-1}x \\ &= i\langle U^{-1}x, e_1\rangle Ue_1 \\ &= i\langle x, Ue_1\rangle Ue_1 \\ &= i\langle x, e_2\rangle e_2 = (T_2 - T_2^*)x,\end{aligned}$$

for each x in $\mathcal{H}_2$. Thus T_2 and T_3 are both quasi-nilpotent operators which leave invariant each subspace in $\mathcal{F}_2$, and they have the same (compact) skew-adjoint part. By the uniqueness clause in Theorem 4.4.2,

$$T_2 = T_3 = UT_1U^{-1}.$$

THEOREM 4.5.3. *Suppose that V is the bounded linear operator, acting on the space $L_2(0,1)$, which is defined by equation (1); and let T be a bounded linear operator which acts on a Hilbert*

space $\mathcal{H}$ *and satisfies conditions (a), (b) and (c) above. Then there is an isometric isomorphism* U *from* $L_2(0,1)$ *onto* $\mathcal{H}$ *such that* $T = UVU^{-1}$.

Proof. This follows immediately from Lemma 4.5.2, by taking $T_1 = V$, $T_2 = T$.

THEOREM 4.5.4. *Suppose that the bounded linear operator* V *acting on the space* $L_2(0,1)$, *and the closed subspaces* L_λ $(0 \leqslant \lambda \leqslant 1)$ *of* $L_2(0,1)$, *are defined by equations (1) and (6). Then, if* M *is a closed subspace of* $L_2(0,1)$ *which is invariant under* V, *there exists a real number* λ *such that* $0 \leqslant \lambda \leqslant 1$ *and* $M = L_\lambda$.

Proof. We have seen, in §4.3.12, that the family $\{L_\lambda : 0 \leqslant \lambda \leqslant 1\}$ is a simple chain $\mathcal{F}_1$ of closed subspaces of $L_2(0,1)$; and each L_λ is invariant under V. Suppose that M is a closed subspace of $L_2(0,1)$ and $V(M) \subseteq M$. The family consisting of the single subspace M is an invariant chain for V, and is therefore contained in a maximal invariant chain $\mathcal{F}_2$, which is simple by Lemma 4.3.2.

We now apply Lemma 4.5.2, with $\mathcal{H}_1 = \mathcal{H}_2 = L_2(0,1)$, $T_1 = T_2 = V$ and $e_1 = e_2 = e$, the unit vector defined by (3). We conclude that there is a unitary operator U on $L_2(0,1)$ for which

$$Ue = e,\ UVU^{-1} = V, \{U(L_\lambda) : 0 \leqslant \lambda \leqslant 1\} = \mathcal{F}_2.$$

These conditions imply that $UV^nU^{-1} = V^n$, and hence that $UV^ne = V^nUe = V^ne$ $(n = 0, 1, 2, \ldots)$. Since finite linear coombinations of the vectors V^ne $(n = 0, 1, 2, \ldots)$ are everywhere dense in $L_2(0,1)$, it follows that $U = I$. Thus

$$M \in \mathcal{F}_2 = \{U(L_\lambda) : 0 \leqslant \lambda \leqslant 1\} = \{L_\lambda : 0 \leqslant \lambda \leqslant 1\}.$$

The following lemma, which is a consequence of Theorem 4.5.3, will be needed in our subsequent work on the convergence of superdiagonal integrals.

LEMMA 4.5.5. *Suppose that* T *is a quasi-nilpotent operator acting on a Hilbert space* $\mathcal{H}$, *and there is a unit vector* f *in* $\mathcal{H}$ *such that*

$$(T-T^*)x = i\langle x, f\rangle f \quad (x \in \mathcal{H}).$$

Then T *is compact, and the non-zero eigenvalues of* $T+T^*$ *are precisely the numbers* $\pm 2/(2n-1)\pi$ $(n = 1, 2, \ldots)$, *each having multiplicity* 1.

Proof. Since Im T is compact, it follows from Theorem 1.8.8 that T is compact. The closed subspace $M = \mathfrak{L}(f, Tf, T^2f, \ldots)$ of $\mathcal{H}$ is invariant under T, so $M^\perp$ is invariant under T^*. If $x \in M^\perp$ then

$$Tx = T^*x + i\langle x, f\rangle f = T^*x \in M^\perp; \tag{11}$$

hence $M^\perp$ is invariant under both T and T^*, and therefore the same is true of M. If $x, y \in M^\perp$ then, by (11),

$$\langle Tx, y\rangle = \langle x, T^*y\rangle = \langle x, Ty\rangle.$$

It follows that the quasi-nilpotent operator obtained by restricting T to $M^\perp$ is self-adjoint, and is therefore zero. We have now proved that

$$T(M) \subseteq M, \qquad T^*(M) \subseteq M, \qquad T(M^\perp) = T^*(M^\perp) = \{0\}. \tag{12}$$

Let T_0 be the compact linear operator obtained by restricting T to M. It follows easily from (12) that T_0^* is the restriction of T^* to M, and that the non-zero eigenvalues of $T_0+T_0^*$ are the same as those of $T+T^*$, with the same multiplicities. However, T_0 satisfies the conditions of Theorem 4.5.3, and is therefore unitarily equivalent

to the operator V defined by equation (1). It now suffices to prove that the non-zero eigenvalues of $V+V^*$ are precisely the numbers $\pm 2/(2n-1)\pi$ ($n = 1, 2, \ldots$), each with multiplicity 1.

For x in $L_2(0,1)$,

$$\begin{aligned}(Vx+V^*x)(s) &= i\int_0^s x(t)dt - i\int_s^1 x(t)dt \\ &= 2i\int_0^s x(t)dt - i\int_0^1 x(t)dt \\ &= 2i\int_0^s x(t)dt - i\langle x, e\rangle,\end{aligned}$$

where e is defined by equation (3). If $\lambda \neq 0$ and $(V+V^*)x = \lambda x$, then

$$2i\int_0^s x(t)dt - i\langle x, e\rangle = \lambda x(s)$$

for almost all s in $[0,1]$. Upon multiplying the last equation by $\lambda^{-1}\exp(-2is/\lambda)$, we deduce that the absolutely continuous function

$$\exp(-2is/\lambda)[\int_0^s x(t)dt - \tfrac{1}{2}\langle x, e\rangle]$$

has derivative zero almost everywhere on $[0,1]$, and is therefore a constant k. Thus

$$\int_0^s x(t)dt = \tfrac{1}{2}\langle x, e\rangle + k\exp(2is/\lambda),$$

and differentiation gives $x(s) = c\exp(2is/\lambda)$ for almost all s in $[0,1]$, where $c = 2ik/\lambda$. It follows that, if λ is a non-zero eigenvalue of $V+V^*$, then every associated eigenvector is a scalar multiple of the vector e_λ defined by $e_\lambda(s) = \exp(2is/\lambda)$; so λ has multiplicity 1. Computation shows that, for any non-zero scalar λ,

$$(V+V^*)e_\lambda = \lambda e_\lambda - \tfrac{1}{2}\lambda[1+\exp(2i/\lambda)]e.$$

Hence λ is an eigenvalue of $V+V^*$ if and only if $\exp(2i/\lambda) = -1$; that is, $\lambda = \pm 2/(2n-1)\pi$ for some positive integer n.

LEMMA 4.5.6. *Suppose that $\mathcal{F}$ is a continuous chain of closed subspaces of a Hilbert space $\mathcal{H}$, K is a self-adjoint operator of finite rank m, the non-zero eigenvalues of K (arranged in order of decreasing magnitude and counted according to their multiplicities) are $\lambda_1, \ldots, \lambda_m$, and (α_n) is the real sequence defined by $\alpha_{2n-1} = \alpha_{2n} = 2/(2n-1)\pi$ ($n = 1, 2, \ldots$). Then the superdiagonal integral*

$$T = 2i\int_{\mathcal{F}} E(M)KE(dM)$$

converges, and its self-adjoint part satisfies

$$||\mathrm{Re}\,T|| \leqslant 2\sum_{j=1}^{m} \alpha_j|\lambda_j|.$$

Proof. Since $M = M_-$ for each M in $\mathcal{F}$, it follows from Theorem 4.4.10 that the superdiagonal integral converges and that T is a Schmidt class operator; by Theorem 4.4.6, Im $T = K$, and T leaves invariant each subspace in $\mathcal{F}$. Thus $T = H+iK$, where H is the self-adjoint Schmidt class operator $\tfrac{1}{2}(T+T^*)$. By Theorem 1.9.2 and equation 1.9(3), there is a unit vector f in $\mathcal{H}$ such that

$$(13) \qquad ||H|| = |\langle Hf, f\rangle|.$$

If B is the operator defined by

$$(14) \qquad Bx = \langle x, f\rangle f \quad (x \in \mathcal{H}),$$

the above reasoning can be applied with B in place of K; in this way we obtain Schmidt class operators A, S such that

$$(15) \qquad A = A^*, \qquad S = A+iB = 2i\int_{\mathcal{F}} E(M)BE(dM).$$

Since the Schmidt class operators S and T leave invariant each subspace $\mathcal{F}$, the same is true of the trace class operator TS.

Proof. The proof is divided into a number of stages.

(a) We consider first the case in which $\mathcal{F}$ is a *continuous* chain. There is an orthonormal system $\{\varphi_j\}$ in $\mathcal{H}$ such that

$$Kx = \sum_{j=1}^{\infty} \lambda_j \langle x, \varphi_j \rangle \varphi_j \quad (x \in \mathcal{H}), \tag{22}$$

and we define approximating operators $K_1, K_2, \ldots$ of finite rank by

$$K_n x = \sum_{j=1}^{n} \lambda_j \langle x, \varphi_j \rangle \varphi_j \quad (x \in \mathcal{H}). \tag{23}$$

By Theorems 4.4.10 and 4.4.6, the superdiagonal integral

$$T_n = 2i \int_{\mathcal{F}} E(M) K_n E(dM) \tag{24}$$

converges, and T_n is a Schmidt class operator which leaves invariant each subspace in $\mathcal{F}$ and has skew-adjoint part K_n. If $m > n \geqslant 1$, then

$$(K_m - K_n)x = \sum_{j=n+1}^{m} \lambda_j \langle x, \varphi_j \rangle \varphi_j \quad (x \in \mathcal{H}) \tag{25}$$

(so the eigenvalues of $K_m - K_n$ are $\lambda_{n+1}, \ldots, \lambda_m$), and

$$T_m - T_n = 2i \int_{\mathcal{F}} E(M)(K_m - K_n)E(dM).$$

By Lemma 4.5.6,

$$||\mathrm{Re}(T_m - T_n)|| \leqslant 2 \sum_{j=1}^{m-n} \alpha_j |\lambda_{n+j}|,$$

where $\alpha_{2r-1} = \alpha_{2r} = 2/(2r-1)\pi$. Furthermore, by (25),

$$||\mathrm{Im}(T_m - T_n)|| = ||K_m - K_n|| = |\lambda_{n+1}|,$$

so

$$||T_m - T_n|| \leqslant |\lambda_{n+1}| + 2 \sum_{j=1}^{m-n} \alpha_j |\lambda_{n+j}|. \tag{26}$$

It follows from (21) that the series

$$\sum_{n=1}^{\infty} \alpha_j |\lambda_j| \tag{27}$$

converges. If N is any positive integer then, since both the sequences (α_j) and $(|\lambda_j|)$ are decreasing,

$$\begin{aligned} ||T_m - T_n|| &\leqslant |\lambda_{n+1}| + 2 \sum_{j=1}^{m-n} \alpha_j |\lambda_{n+j}| \\ &\leqslant |\lambda_{n+1}| + 2 \sum_{j=1}^{N} \alpha_j |\lambda_{n+1}| + 2 \sum_{j=N+1}^{\infty} \alpha_j |\lambda_j|, \\ &= |\lambda_{n+1}| [1 + 2 \sum_{j=1}^{N} \alpha_j] + 2 \sum_{j=N+1}^{\infty} \alpha_j |\lambda_j|. \end{aligned}$$

Given a positive real number ϵ, we can choose N such that

$$2 \sum_{j=N+1}^{\infty} \alpha_j |\lambda_j| < \tfrac{1}{2}\epsilon,$$

since the series (27) converges; then, since $\lambda_n \to 0$ as $n \to \infty$, we can choose an integer N_0 such that

$$|\lambda_{n+1}| \, [1 + 2 \sum_{j=1}^{N} \alpha_j] < \tfrac{1}{2}\epsilon \quad \text{whenever } n > N_0.$$

It follows that $||T_m - T_n|| < \epsilon$ whenever $m > n > N_0$. Thus (T_n) is a Cauchy sequence in $\mathcal{B}(\mathcal{H})$, and so converges in norm to an operator T in $\mathcal{B}(\mathcal{H})$. Since T_n is compact and leaves invariant each subspace in $\mathcal{F}$, the same is true of T, so T is quasi-nilpotent by Corollary 4.3.11. Furthermore

$$\mathrm{Im}\, T = \lim_{n \to \infty} \mathrm{Im}(T_n) = \lim_{n \to \infty} K_n = K.$$

By Theorem 4.4.6, the superdiagonal integral of K relative to $\mathcal{F}$ converges, and

$$2i\int_{\mathcal{F}} E(M)KE(dM) = T.$$

In view of (23) and (24), it follows from Lemma 4.5.6 that

$$\|\mathrm{Re}\ T_n\| \leqslant 2\sum_{j=1}^{n} a_j|\lambda_j|$$
$$\leqslant 2\sum_{j=1}^{\infty} a_j|\lambda_j|,$$

$$\|\mathrm{Re}\ T_n\| \leqslant \frac{4}{\pi}\sum_{n=1}^{\infty} \frac{|\lambda_{2n-1}|+|\lambda_{2n}|}{2n-1};$$

we obtain the same inequality for T be taking limits as $n \to \infty$. This completes the proof of the theorem when $\mathcal{F}$ is a continuous chain.

(b) We now consider the case in which $\mathcal{F}$ is a simple (but not necessarily continuous) chain. Our first step is to embed $\mathcal{H}$ in a larger Hilbert space $\mathcal{H}_0$, and to construct from $\mathcal{F}$ a continuous chain $\mathcal{F}_0$ of closed subspaces of $\mathcal{H}_0$. For this purpose we assign, to each subspace M in $\mathcal{F}$ such that $M \neq M_-$, a separable infinite-dimensional Hilbert space $\mathcal{H}_M$; and we form the Hilbert direct sums

(28) $$\mathcal{H}_1 = \Sigma\oplus\{\mathcal{H}_M : M \in \mathcal{F}, M \neq M_-\}, \qquad \mathcal{H}_0 = \mathcal{H}\oplus\mathcal{H}_1.$$

If $M \in \mathcal{F}_0$ and $M \neq M_-$, then $P(M)(\mathcal{H})$ is a 1-dimensional subspace of $\mathcal{H}$, so

(29) $$\mathcal{K}_M = P(M)(\mathcal{H})\oplus\mathcal{H}_M$$

is a separable infinite-dimensional Hilbert space, and can be embedded canonically as a subspace of $\mathcal{H}_0$ $(= \mathcal{H}\oplus\mathcal{H}_1)$. It follows from the discussion at the end of §4.3.12 that there is a continuous chain $\mathcal{F}_M$ of closed subspaces of $\mathcal{K}_M$. When $M \in \mathcal{F}$ and $M = M_-$, we define $\mathcal{H}_M = \mathcal{K}_M = \{0\}$, and denote by $\mathcal{F}_M$ the trivial chain consisting of $\mathcal{K}_M$ only. We shall identify each of the Hilbert spaces $\mathcal{H}$, $\mathcal{H}_1$, $\mathcal{H}_M$, $\mathcal{K}_M$ with its canonical image in $\mathcal{H}_0$.

If $M \in \mathcal{F}$, the subspaces M_- and $P(M)(\mathcal{H})$ of $\mathcal{H}$ are mutually orthogonal, since $P(M) = E(M)-E(M_-)$. From this, together with (28) and (29), it follows that the subspaces

$$M_-, \qquad \mathcal{K}_M, \qquad \mathcal{H}_L \quad (L \in \mathcal{F},\ L \subseteq M_-)$$

are pairwise orthogonal. Given any M in $\mathcal{F}$ and N in $\mathcal{F}_M$, $N \subseteq \mathcal{K}_M$ and we can form the subspace

$$X(M, N) = M_-\oplus N\oplus\Sigma\oplus\{\mathcal{H}_L : L \in \mathcal{F}, L \subset M_-\}$$

of $\mathcal{H}_0$. As N increases continuously from $\{0\}$ to $\mathcal{K}_M$, $X(M, N)$ increases continuously from

$$M_-\oplus\Sigma\oplus\{\mathcal{H}_L : L \in \mathcal{F}, L \subseteq M_-\}$$

to $M\oplus\Sigma\oplus\{\mathcal{H}_L : L \in \mathcal{F}, L \subseteq M\}$.

With this is mind, it is not difficult to show that

$$\mathcal{F}_0 = \{X(M, N) : M \in \mathcal{F}, N \in \mathcal{F}_M\}$$

is a continuous chain of closed subspaces of $\mathcal{H}_0$; the details of this verification are left as an exercise for the reader.

(c) Let K_0 be the compact self-adjoint operator on $\mathcal{H}_0$ defined by

(30) $$K_0(x+x_1) = Kx \quad (x \in \mathcal{H},\ x_1 \in \mathcal{H}_1).$$

Since K_0 has the same non-zero eigenvalues as K, and $\mathcal{F}_0$ is a continuous chain, it follows from part (a) of this proof that the superdiagonal integral

$$T_0 = 2i\int_{\mathcal{F}_0} E(M)K_0E(dM)$$

converges: furthermore

(31) $$||\tfrac{1}{2}(T_0+T_0^*)|| < \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{|\lambda_{2n-1}|+|\lambda_{2n}|}{2n-1}.$$

Of course, T_0 is a compact quasi-nilpotent operator with skew-adjoint part K_0, and leaves invariant each subspace in $\mathcal{F}_0$.

(d) We shall prove that

(32) $$T_0(\mathcal{H}_1) = \{0\}.$$

For this purpose, suppose that $M \in \mathcal{F}$ and $M \neq M_-$. Then the subspaces

$$X_1 = X(M, \{0\}) = M_- \oplus \Sigma\oplus \{\mathcal{H}_L : L \in \mathcal{F}, L \subseteq M_-\},$$

$$X_2 = X(M, \mathcal{K}_M) = M \oplus \Sigma\oplus \{\mathcal{H}_L : L \in \mathcal{F}, L \subseteq M\}$$

of $\mathcal{H}_0$ belong to the chain $\mathcal{F}_0$, and

(33) $$X_1 \subseteq X_2, \qquad X_2 \cap X_1^{\perp} = \mathcal{K}_M = P(M)(\mathcal{H}) \oplus \mathcal{H}_M.$$

Let E_j be the projection from $\mathcal{H}_0$ onto X_j, and consider the operator $(E_2-E_1)K_0(E_2-E_1)$. Since

(34) $$(E_2-E_1)(\mathcal{H}_0) = X_2 \cap X_1^- = P(M)(\mathcal{H}) \oplus \mathcal{H}_M,$$

and $\mathcal{H}_M \subset \mathcal{H}_1$, it follows from (30) that

$$K_0(E_2-E_1)(x+x_1) = KP(M)x \quad (x \in \mathcal{H},\ x_1 \in \mathcal{H}_1).$$

Since the right-hand side of this last equation is in $\mathcal{H}$ and so orthogonal to $\mathcal{H}_M$, (34) now implies that

$$(E_2-E_1)K_0(E_2-E_1)(x+x_1) = P(M)KP(M)x = 0.$$

Thus $(E_2-E_1)K_0(E_2-E_1) = 0$, and

$$\begin{aligned}\mathrm{Im}(E_2-E_1)T_0(E_2-E_1) &= (E_2-E_1)(\mathrm{Im}\ T_0)(E_2-E_1)\\ &= (E_2-E_1)K_0(E_2-E_1) = 0;\end{aligned}$$

so the operator $(E_2-E_1)T_0(E_2-E_1)$ is self-adjoint.

Given any bounded linear operators A_1 and A_2 on $\mathcal{H}_0$, which leave invariant the subspaces X_1 and X_2, we have $A_iE_j = E_jA_iE_j$ (equivalently, $(I-E_j)A_i = (I-E_j)A_i(I-E_j)$) for $i,j = 1,2$; thus

$$\begin{aligned}&(E_2-E_1)A_1A_2(E_2-E_1)\\ &\quad= (E_2-E_1)(I-E_1)A_1A_2E_2(E_2-E_1)\\ &\quad= (E_2-E_1)(I-E_1)A_1(I-E_1)E_2A_2E_2(I-E_2)\\ &\quad= (E_2-E_1)A_1(E_2-E_1)A_2(E_2-E_1).\end{aligned}$$

By applying this with A_1 and A_2 powers of T_0, we prove inductively that

$$[(E_2-E_1)T_0(E_2-E_1)]^n = (E_2-E_1)T_0^n(E_2-E_1)$$

$(n = 1, 2, \ldots)$. Thus

$$||[(E_2-E_1)T_0(E_2-E_1)]^n|| \leqslant ||T_0^n||,$$

and so $(E_2-E_1)T_0(E_2-E_1)$ is quasi-nilpotent; since it is also self-adjoint,

(35) $$(E_2-E_1)T_0(E_2-E_1) = 0.$$

If $x \in \mathcal{H}_M(\subset \mathcal{H}_1)$ then, by (30),

$$(\mathrm{Im}\ T_0)x = K_0x = 0,$$

so $T_0x = T_0^*x$; furthermore,

$$x \in X_2 \cap X_1^{\perp} = (E_2-E_1)(\mathcal{H}_0).$$

Now T_0 leaves X_1 and X_2 invariant, thus T_0^* leaves $X_1^{\perp}$ invariant, and so

$$T_0x \in X_2, \qquad T_0x = T_0^*x \in X_1^{\perp}.$$

Thus $T_0x \in X_2 \cap X_1^{\perp}$, and so

$$T_0x = (E_2-E_1)T_0(E_2-E_1)x = 0$$

by (35). This shows that $T_0(\mathcal{H}_M) = \{0\}$ whenever $M \in \mathfrak{F}$ and $M \neq M_-$, and it follows from (28) that $T_0(\mathcal{H}_1) = \{0\}$, proving (32).

(e) By (30) and (32),

$$T_0(\mathcal{H}_1) = \{0\},\ (\operatorname{Im} T_0)(\mathcal{H}_1) = K_0(\mathcal{H}_1) = \{0\},$$

so $T_0^*(\mathcal{H}_1) = \{0\}$. Since $\mathcal{H}$ is the orthogonal complement of $\mathcal{H}_1$ in $\mathcal{H}_0$, it follows that $T_0(\mathcal{H}_0) \subseteq \mathcal{H}$. This, with (32), shows that there is a bounded linear operator T on $\mathcal{H}$ such that

$$T_0(x+x_1) = Tx \quad (x \in \mathcal{H},\ x_1 \in \mathcal{H}_1),$$

whence also

$$T_0^*(x+x_1) = T^*x \quad (x \in \mathcal{H},\ x_1 \in \mathcal{H}_1).$$

It is clear that T is quasi-nilpotent, $\operatorname{Im} T = K$ and, by (31),

$$\|\tfrac{1}{2}(T+T^*)\| = \|\tfrac{1}{2}(T_0+T_0^*)\| \leqslant \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{|\lambda_{2n-1}|+|\lambda_{2n}|}{2n-1}.$$

Given M in $\mathfrak{F}$, $M = X(M, \mathcal{K}_M) \cap \mathcal{H}$ (as a subspace of $\mathcal{H}_0$); since $X(M, \mathcal{K}_M)$ and $\mathcal{H}$ are both invariant under T_0, so is M, and thus M (as a subspace of $\mathcal{H}$) is invariant under T. It now follows from Theorem 4.4.6 that the superdiagonal integral

$$2i\int_{\mathfrak{F}} E(M)KE(dM)$$

converges, and that its value is T. This completes the proof of Theorem 4.5.7.

The hypotheses of Theorem 4.5.7 are satisfied if K is a self-adjoint element of a von Neumann-Schatten class $\mathcal{C}_p$, where $1 < p < \infty$, and $P(M)KP(M) = 0$ whenever $M \in \mathfrak{F}$, $M \neq M_-$; for the index q conjugate to p satisfies $q > 1$, and

$$\sum n^{-1}|\lambda_n| \leqslant (\sum n^{-q})^{1/q}(\sum |\lambda_n|^p)^{1/p} < \infty.$$

A good deal more is known [45, 20]; if K satisfies the above conditions and

$$T = 2i\int_{\mathfrak{F}} E(M)KE(dM)$$

then $T \in \mathcal{C}_p$; furthermore, the mapping $K \to T$ is bounded relative to the norm on $\mathcal{C}_p$. These results fail when $p = 1$ (consider the operator V defined by (1)).

4.6. *An alternative treatment of the integration operator in* $L_2(0,1)$

The proofs given in §4.5 for Theorems 4.5.3 and 4.5.4 are heavily dependent on the theory of superdiagonal representation of compact linear operators; in this section, we give a more elementary treatment. Once again, we assume throughout that T is a bounded linear operator acting on a Hilbert space $\mathcal{H}$, and

(a) T is quasi-nilpotent,

(b) there is a unit vector f in $\mathcal{H}$ such that

$$(T-T^*)x = i\langle x, f\rangle f \quad (x \in \mathcal{H}),$$

(c) $\mathfrak{L}(f, Tf, T^2f, \ldots) = \mathcal{H}$.

Condition (b) may be expressed in the form

$$T-T^* = iK, \tag{1}$$

where K is the self-adjoint operator given by

$$Kx = \langle x, f\rangle f \quad (x \in \mathcal{H}). \tag{2}$$

We shall consider the entire function φ defined by

(3) $$\varphi(\lambda) = 1-i\lambda\langle(I+\lambda T)^{-1}f,\ f\rangle.$$

LEMMA 4.6.1. *For all complex numbers λ and μ,*

$$\varphi(\lambda)\overline{\varphi(\mu)} = 1+\frac{\lambda-\overline{\mu}}{i}\langle(I+\lambda T)^{-1}f,\ (I+\mu T)^{-1}f\rangle.$$

Proof. Since

$$\begin{aligned}(\lambda-\overline{\mu})&(I+\overline{\mu}T^*)^{-1}(I+\lambda T)^{-1}\\ &= (I+\overline{\mu}T^*)^{-1}[\lambda(I+\overline{\mu}T^*)-\overline{\mu}(I+\lambda T)+\lambda\overline{\mu}(T-T^*)](I+\lambda T)^{-1}\\ &= \lambda(I+\lambda T)^{-1}-\overline{\mu}(I+\overline{\mu}T^*)^{-1}+i\lambda\overline{\mu}(I+\overline{\mu}T^*)^{-1}K(I+\lambda T)^{-1}\end{aligned}$$

it follows that

$$\begin{aligned}(\lambda-\overline{\mu})&\langle(I+\lambda T)^{-1}f,\ (I+\mu T)^{-1}f\rangle\\ &= \langle(\lambda-\overline{\mu})(I+\overline{\mu}T^*)^{-1}(I+\lambda T)^{-1}f,\ f\rangle\\ &= \lambda\langle(I+\lambda T)^{-1}f,\ f\rangle-\overline{\mu}\langle(I+\overline{\mu}T^*)^{-1}f,\ f\rangle+\\ &\qquad +i\lambda\overline{\mu}\langle(I+\overline{\mu}T^*)^{-1}K(I+\lambda T)^{-1}f,\ f\rangle\end{aligned}$$

Now

$$K(I+\lambda T)^{-1}f = \langle(I+\lambda T)^{-1}f,\ f\rangle f,$$

so

$$\langle(I+\overline{\mu}T^*)^{-1}K(I+\lambda T)^{-1}f,\ f\rangle = \langle(I+\lambda T)^{-1}f,\ f\rangle\langle(I+\overline{\mu}T^*)^{-1}f,\ f\rangle.$$

Thus

$$\begin{aligned}(\lambda-\overline{\mu})&\langle(I+\lambda T)^{-1}f,\ (I+\mu T)^{-1}f\rangle\\ &= \lambda\langle(I+\lambda T)^{-1}f,\ f\rangle-\overline{\mu}\langle(I+\overline{\mu}T^*)^{-1}f,\ f\rangle+\\ &\qquad +i\lambda\overline{\mu}\langle(I+\lambda T)^{-1}f,\ f\rangle\langle(I+\overline{\mu}T^*)^{-1}f,\ f\rangle,\end{aligned}$$

and

$$\begin{aligned}1+&\frac{\lambda-\overline{\mu}}{i}\langle(I+\lambda T)^{-1}f,\ (I+\mu T)^{-1}f\rangle\\ &= 1-i\lambda\langle(I+\lambda T)^{-1}f,\ f\rangle+i\overline{\mu}\langle(I+\overline{\mu}T^*)^{-1}f,\ f\rangle+\\ &\qquad +\lambda\overline{\mu}\langle(I+\lambda T)^{-1}f,\ f\rangle\langle(I+\overline{\mu}T^*)^{-1}f,\ f\rangle\\ &= [1-i\lambda\langle(I+\lambda T)^{-1}f,\ f\rangle][1+i\overline{\mu}\langle(I+\overline{\mu}T^*)^{-1}f,\ f\rangle]\\ &= \varphi(\lambda)\ \overline{\varphi(\mu)}.\end{aligned}$$

LEMMA 4.6.2. *For all complex numbers λ, $\varphi(\lambda) = \exp(-i\lambda)$.*

Proof. By Lemma 4.6.1,

$$|\varphi(\lambda)|^2 = 1+2(\operatorname{Im}\lambda)||(I+\lambda T)^{-1}f||^2,$$

and

$$\varphi(\lambda)\overline{\varphi(\mu)} = 1$$

if $\mu = \overline{\lambda}$. Furthermore, (3) can be written in the form

(4) $$\varphi(\lambda) = 1+\sum_{n=1}^{\infty}(-1)^n i\langle T^{n-1}f,\ f\rangle\lambda^n.$$

It follows that φ is an entire function which does not take the value zero, $\varphi(0) = 1$, $|\varphi(\lambda)| = 1$ when λ is real and $|\varphi(\lambda)| > 1$ when $\operatorname{Im}\lambda > 0$. By Theorem 1.4.3 there is a positive real number a such that $\varphi(\lambda) = \exp(-ia\lambda)$ $(\lambda \in \mathbf{C})$. Since the coefficient of λ on the right-hand side of (4) is $-i$, it follows that $a = 1$.

Alternative proof of Theorem 4.5.3. We have seen that, if the unit vector e in $L_2(0,1)$ and the integration operator V are defined by equations 4.5(3) and 4.5(1), then V satisfies conditions (a), (b) and (c) above, with f replaced by e. Suppose that T is any other operator which satisfies these conditions. It follows from Lemmas 4.6.1 and 4.6.2 that

$$1+\frac{\lambda-\bar{\mu}}{i}\langle (I+\lambda T)^{-1}f,\ (I+\mu T)^{-1}f\rangle$$

$$= \exp i(\bar{\mu}-\lambda) = 1+\frac{\lambda-\bar{\mu}}{i}\langle (I+\lambda V)^{-1}e,\ (I+\mu V)^{-1}e\rangle,$$

whence

$$\langle (I+\lambda T)^{-1}f,\ (I+\mu T)^{-1}f\rangle = \langle (I+\lambda V)^{-1}e,\ (I+\mu V)^{-1}e\rangle,$$

for all complex λ and μ. By expanding both sides of this last equation in powers of λ and $\bar{\mu}$, and comparing coefficients, we obtain

$$\langle T^m f,\ T^n f\rangle = \langle V^m e,\ V^n e\rangle \quad (m, n = 0, 1, 2, \ldots).$$

By linearity of the inner products, if p and q are polynomials with complex coefficients, then

$$\langle p(T)f,\ q(T)f\rangle = \langle p(V)e,\ q(V)e \quad ;$$

in particular

(5) $$\|p(T)f\| = \|p(V)e\|.$$

Since T and V both satisfy condition (c), the set of all vectors $p(V)e$ is an everywhere dense subspace X of $L_2(0,1)$, and vectors of the form $p(T)f$ are dense in the Hilbert space $\mathcal{H}$ on which T acts. If $x \in X$ and x is represented in two ways, $x = p(V)e = q(V)e$, then

$$\|p(T)f-q(T)f\| = \|p(V)e-q(V)e\| = 0,$$

by (5), and so $p(T)f = q(T)f$. It follows that the equation

(6) $$Up(V)e = p(T)f$$

defines (unambiguously) an isometric isomorphism from X onto an everywhere dense subspace of $\mathcal{H}$; this extends by continuity to an isometric isomorphism, which we denote by the same symbol U, from $L_2(0,1)$ onto $\mathcal{H}$. For $n = 0, 1, 2, \ldots,$

$$\begin{aligned} UVV^n e &= UV^{n+1}e \\ &= T^{n+1}f \\ &= TT^n f \\ &= TUV^n e. \end{aligned}$$

By linearity and continuity, $UVx = TUx$ for each x in $\mathfrak{L}(e, Ve, V^2e, \ldots)$ $(= L_2(0,1))$; so $UV = TU$, $T = UVU^{-1}$. This completes the alternative proof of Theorem 4.5.3.

In the following lemmas, which lead up to an alternative proof of Theorem 4.5.4, the unit vector e in $L_2(0,1)$ and the integration operator V are defined as in §4.5, and $\mathcal{H}_0$ is a closed subspace of $L_2(0,1)$ which is invariant under V. The restriction of V to $\mathcal{H}_0$ is denoted by V_0, and P is the projection from $L_2(0,1)$ onto $\mathcal{H}_0$.

LEMMA 4.6.3. $\mathcal{H}_0 = \mathfrak{L}(Pe, VPe, V^2Pe, \ldots)$.

Proof. If $\mathcal{H}_1 = \mathfrak{L}(Pe, VPe, V^2Pe, \ldots)$, then $\mathcal{H}_1 \subseteq \mathcal{H}_0$; in order to prove that $\mathcal{H}_1 = \mathcal{H}_0$, it is sufficient to show that $\mathcal{H}_0 \cap \mathcal{H}_1^{\perp} = \{0\}$. Clearly $V(\mathcal{H}_1) \subset \mathcal{H}_1$, so

$$V^*(\mathcal{H}_1^{\perp}) \subseteq \mathcal{H}_1^{\perp}.$$

Suppose that $y \in \mathcal{H}_0 \cap \mathcal{H}_1^{\perp}$. Then

$$\langle y,\ e\rangle = \langle Py,\ e\rangle = \langle y,\ Pe\rangle = 0,$$

and $Vy - V^*y = i\langle y,\ e\rangle e = 0$. Thus $Vy = V^*y$,

$$Vy \in V(\mathcal{H}_0) \subseteq \mathcal{H}_0, \qquad V^*y \in V^*(\mathcal{H}_1^{\perp}) \subseteq \mathcal{H}_1^{\perp},$$

and so $Vy = V^*y \in \mathcal{H}_0 \cap \mathcal{H}_1^{\perp}$. We have now proved that

(7) $$\begin{cases} \langle y,\ e\rangle = 0, \\ Vy = V^*y \in \mathcal{H}_0 \cap \mathcal{H}_1^{\perp}, \text{ whenever } y \in \mathcal{H}_0 \cap \mathcal{H}_1^{\perp}. \end{cases}$$

Given any x in $\mathcal{H}_0 \cap \mathcal{H}_1^{\perp}$, it follows from (7) that $(V^*)^n x \in \mathcal{H}_0 \cap \mathcal{H}_1^{\perp}$, and hence that $\langle (V^*)^n x, e\rangle = 0$, for $n = 0, 1, 2, \ldots$ Thus $\langle x, V^n e\rangle = 0$ $(n = 0, 1, 2, \ldots)$ and, since $\mathfrak{L}(e, Ve, V^2 e, \ldots) = L_2(0,1)$, $x = 0$. This shows that $\mathcal{H}_0 \cap \mathcal{H}_1^{\perp} = \{0\}$, as required.

LEMMA 4.6.4. *For all complex numbers* λ,

$$1 - i\lambda \langle (I+\lambda V)^{-1} Pe, e\rangle = \exp(-i\|Pe\|^2 \lambda).$$

Proof. If $\mathcal{H}_0 = \{0\}$ then $P = 0$ and the result is apparent. We suppose $\mathcal{H}_0 \neq \{0\}$, so that $Pe \neq 0$ by Lemma 4.6.3, and we define $\sigma = \|Pe\|$ (> 0). If $x, y \in \mathcal{H}_0$, then

$$\begin{aligned}\langle x, V_0^* y\rangle = \langle V_0 x, y\rangle &= \langle Vx, y\rangle \\ &= \langle x, V^* y\rangle \\ &= \langle Px, V^* y\rangle = \langle x, PV^* y\rangle.\end{aligned}$$

Since $V_0^* y, PV^* y \in \mathcal{H}_0$, it follows that $V_0^* y = PV^* y$. Also, $V_0 y = Vy = PVy$, because $Vy \in \mathcal{H}_0$. Thus

$$\begin{aligned}V_0 y - V_0^* y &= P(Vy - V^* y) \\ &= i\langle y, e\rangle Pe \\ &= i\langle Py, e\rangle Pe \\ &= i\langle y, Pe\rangle Pe \quad (y \in \mathcal{H}_0).\end{aligned}$$

If f denotes the unit vector $\sigma^{-1} Pe$ in $\mathcal{H}_0$, and $T = \sigma^{-2} V_0$, then the last equation asserts that

$$Ty - T^* y = i\langle y, f\rangle f \quad (y \in \mathcal{H}_0).$$

Now T is quasi-nilpotent, and it follows from Lemma 4.6.3 that $\mathfrak{L}(f, Tf, T^2 f, \ldots) = \mathcal{H}_0$. Hence T satisfies conditions (a), (b), (c) above and, by Lemma 4.6.2,

$$1 - i\lambda\langle (I+\lambda T)^{-1} f, f\rangle = \exp(-i\lambda)$$

for all complex λ. Thus

$$\begin{aligned}\exp(-i\|Pe\|^2\lambda) &= \exp(-i\sigma^2\lambda) \\ &= 1 - i\sigma^2\lambda\langle (I+\sigma^2\lambda T)^{-1} f, f\rangle \\ &= 1 - i\lambda\langle (I+\lambda V_0)^{-1} Pe, Pe\rangle \\ &= 1 - i\lambda\langle P(I+\lambda V_0)^{-1} Pe, e\rangle.\end{aligned}$$

Since $(I+\lambda V_0)^{-1} Pe \in \mathcal{H}_0$,

$$\begin{aligned}P(I+\lambda V_0)^{-1} Pe &= (I+\lambda V_0)^{-1} Pe \\ &= \sum_{n=0}^{\infty} (-\lambda V_0)^n Pe \\ &= \sum_{n=0}^{\infty} (-\lambda V)^n Pe = (I+\lambda V)^{-1} Pe,\end{aligned}$$

and so

$$\exp(-i\|Pe\|^2\lambda) = 1 - i\lambda\langle (I+\lambda V)^{-1} Pe, e\rangle.$$

LEMMA 4.6.5. *Suppose that* P_j *is the projection from* $L_2(0,1)$ *onto a closed subspace* $\mathcal{H}_j$ *which is invariant under* V $(j = 1,2)$. *If* $\|P_1 e\| = \|P_2 e\|$, *then* $\mathcal{H}_1 = \mathcal{H}_2$.

Proof. If $\|P_1 e\| = \|P_2 e\| = \sigma$ then, by Lemma 4.6.4,

$$1 - i\lambda\langle (I+\lambda V)^{-1} P_1 e, e\rangle = 1 - i\lambda\langle (I+\lambda V)^{-1} P_2 e, e\rangle \, (= \exp(-i\sigma^2\lambda))$$

for all complex λ. By expanding both sides of this equation in powers of λ and comparing coefficients, we obtain $\langle V^n P_1 e, e\rangle = \langle V^n P_2 e, e\rangle$ $(n = 0, 1, 2, \ldots)$. Thus

$$\langle P_1 e - P_2 e, (V^*)^n e\rangle = 0 \quad (n = 0, 1, 2, \ldots). \tag{8}$$

A straightforward computation based on equation 4.5(2) shows that

$$(V^*)^n e = (n!)^{-1}(-i)^n e_n,$$

where the vector e_n in $L_2(0,1)$ is given by $e_n(s) = (1-s)^n$. Thus the set of finite linear combinations of vectors $e, V^*e, (V^*)^2 e, \ldots$ consists of all polynomials, and is therefore dense in $L_2(0,1)$. This, together with (8), shows that $P_1 e = P_2 e$. By Lemma 4.6.3, $\mathcal{H}_j = \mathfrak{L}(P_j e, VP_j e, V^2 P_j e \ldots)$, so $\mathcal{H}_1 = \mathcal{H}_2$.

Alternative proof of Theorem 4.5.4. Suppose that $\mathcal{H}_1$ is a closed subspace of $L_2(0,1)$ which is invariant under V, and that P_1 is the corresponding projection. Let $\sigma = ||P_1 e||$, so that $0 \leqslant \sigma \leqslant 1$, and define a closed subspace $\mathcal{H}_2$ of $L_2(0,1)$ by

$$\mathcal{H}_2 = \{x \in L_2(0,1) : x(s) = 0 \text{ almost everywhere on } [0,1-\sigma^2]\}.$$

Thus $\mathcal{H}_2$ is invariant under V and, in the notation of equation 4.5(6), $\mathcal{H}_2 = L_\lambda$ with $\lambda = 1-\sigma^2$. The projection P_2 from $L_2(0,1)$ onto $\mathcal{H}_2$ is given by

$$(P_2 x)(s) = \begin{cases} 0 & 0 \leqslant s < 1-\sigma^2, \\ x(s) & 1-\sigma^2 \leqslant s < 1, \end{cases}$$

so

$$||P_2 x||^2 = \int_{1-\sigma^2}^{1} |x(s)|^2 ds \quad (x \in L_2(0,1)).$$

In particular, $||P_2 e|| = \sigma = ||P_1 e||$; by Lemma 4.6.5, $\mathcal{H}_1 = \mathcal{H}_2 = L_\lambda$, where $\lambda = 1-\sigma^2$.

Bibliography

[1] ARONSZAJN, N. and SMITH, K.T. 'Invariant subspaces of completely continuous operators', *Ann. Math.*, **60** 345–350 (1954)

[2] ARVESON, W.B. and FELDMAN, J. 'A note on invariant subspaces', *Mich. math. J.*, **15** 61–64 (1968)

[3] BERNSTEIN, A.R. 'Invariant subspaces of polynomially compact operators on Banach space', *Pacif. J. Math.* **21** 445–464 (1967)

[4] BERNSTEIN, A.R. and ROBINSON, A. 'Solution of an invariant subspace problem of K.T. Smith and P.R. Halmos', *Pacif. J. Math.*, **16** 421–431 (1966)

[5] BRODSKIĬ, M.S. 'On a problem of I.M. Gel'fand (Russian), *Usp. mat. Nauk,* **12** No. 2 (74), 129–132 (1957)

[6] BRODSKIĬ, M.S. 'On the triangular representation of completely continuous operators with one-point spectra' (Russian), *Usp. mat. Nauk,* **16** No. 1 (97), 135–141 (1961)

[7] BRODSKIĬ, M.S. 'An abstract triangular representation of completely continuous operators with a point spectrum' (Russian), 'Funkcional'nyĭ Analiz i ego Primenenie' *(Trudy 5 Konf. po Funkcional'nomu i ego Primeneniju)*, 25–28. Izdat. Akad. Nauk Azerbaĭdžan. SSSR, Baku, (1961)

[8] BRODSKIĬ, M.S. 'On the unicellularity of the integration operator and a theorem of Titchmarsh' (Russian), *Usp. mat. Nauk,* **20** No. 5 (125), 189–192 (1965)

[9] BRODSKIĬ, M.S., GOHBERG, I.C., KREĬN, M.G. and MACAEV, V.I. 'Some new investigations in the theory of non-self-adjoint operators' (Russian), *Proc. Fourth All-Union Math. Congr. (Leningrad, 1961)* (Russian), Vol. II, 261-271. Izdat. 'Nauka', Leningrad, (1964)

[10] BRODSKIĬ, M.S. and KISILEVSKIĬ, G.E. 'Criterion for unicellularity of dissipative Volterra operators with nuclear imaginary components' (Russian), *Izv. Akad. Nauk SSSR*, Ser. Mat., **30** 1212–1228 (1966)

[11] CARLEMAN, T. 'Zur Theorie der linearen Integralgleichungen', *Math. Z.*, **9** 196–217 (1921)

[12] COPSON, E.T. *An introduction to the theory of functions of a complex variable* (Oxford 1935)

[13] DECKARD, D., DOUGLAS, R.G. and PEARCY, C. 'On invariant subspaces of quasitriangular operators', *Am. J. Math.*, **91** 637–647 (1969)

[14] DIXMIER, J. 'Les opérateurs permutables à l'opérateur intégral', *Port. math.*, **8** 73–84 (1949)

[15] DONOGHUE, W.F. 'The lattice of invariant subspaces of a completely continuous quasi-nilpotent transformation', *Pacif. J. Math.*, **7** 1031–1035 (1957)

[16] DUNFORD, N. and SCHWARTZ, J.T. *Linear operators*, Part II, (Interscience, New York, 1963)

[17] GILLESPIE, T.A. 'An invariant subspace theorem of J. Feldman', *Pacif. J. Math.*, **26** 67–72 (1968)

[18] GOHBERG, I.C. and KREĬN, M.G. 'Completely continuous operators with spectrum concentrated at zero' (Russian), *Dokl. Akad. Nauk SSSR*, **128** 227–230 (1959)

[19] GOHBERG, I.C. and KREĬN, M.G. 'On the theory of triangular representations of non-self-adjoint operators' (Russian), *Dokl. Akad. Nauk SSSR*, **137** 1034–1037 (1961) English translation: *Soviet Math. Dokl.*, 2^1 392–395 (1961)

[20] GOHBERG, I.C. and KREĬN, M.G. 'Volterra operators whose imaginary component belongs to a given class' (Russian), *Dokl. Akad. Nauk SSSR*, **139** 779–782 (1961) English translation: *Soviet Math. Dokl.*, 2^2 983–986 (1961)

[21] GOHBERG, I.C. and KREĬN, M.G. *Introduction to the theory of linear non-self-adjoint operators* (Russian), Izdat. 'Nauka', (Moscow, 1965)
English translation: *Translations of Mathematical Monographs*, volume 18, American Math. Soc. (1968)

[22] GOHBERG, I.C. and KREĬN, M.G. *Theory of Volterra operators in Hilbert space and its applications* (Russian), Izdat. 'Nauka', (Moscow, 1967)
English translation: *Translations of Mathematical Monographs*, volume 24, American Math. Soc. (1968)

[23] GOL'DFAIN, I.A. 'Sur une classe d'équations intégrales linéaires', *Uchen. Zap. Mosk. gos. Univ.*, **100** *Matematika*, tom 1, 104–112 (1946)
English translation: *Am. math. Soc. Transl. (2)*, **10** 283–290 (1958)

[24] GROTHENDIECK, A. 'La théorie de Fredholm', *Bull. Soc. math. Fr.*, **84** 319–384 (1956)

[25] HALMOS, P.R. *Measure theory*, (Van Nostrand, 1950)

[26] HALMOS, P.R. *Finite-dimensional vector spaces*, (Van Nostrand, *1958)*

[27] HALMOS, P.R. 'Invariant subspaces of polynomially compact operators', *Pacif. J. Math.*, **16** 433–437 (1966)

[28] HSU, N.H. 'Invariant subspaces of polynomially compact operators in Banach spaces', *Yokohama math. J.*, **15** 11–15 (1967)

[29] KALISCH, G.K. 'On similarity, reducing manifolds, and unitary equivalence of certain Volterra operators', *Ann. Math.*, **66** 481–494 (1957)

[30] KALISCH, G.K. 'A functional analysis proof of Titchmarsh's theorem on convolution', *J. math. Analysis. Applic.*, **5** 176–183 (1962)

[31] KALISCH, G.K. 'Direct proofs of spectral representation theorems', *J. math. Analysis Applic.*, **8** 351–363 (1964)

[32] KELDYŠ, M.V. and LIDSKIĬ, V.B. 'On the spectral theory of non-self-adjoint operators' (Russian), *Proc. Fourth All-Union Math. Congr. (Lengingrad, 1961)* (Russian), Vol. I, 101–120. Izdat. Akad, Nauk SSSR, Leningrad, (1963)

[33] KELLEY, J.L. *General topology*, (Van Nostrand, 1955)

[34] KREĬN, M.G. 'Criteria for completeness of the system of root vectors of a dissipative operator' (Russian), *Usp. mat. Nauk,* **14** No. 3 (87), 145–152 (1959)
English translation: *Am. math. Soc. Transl. (2),* **26** 221–229 (1963)

[35] KREĬN, M.G. 'A contribution to the theory of linear non-self-adjoint operators' (Russian), *Dokl. Akad. Nauk SSSR,* **130** 254–256 (1960)
English translation: *Soviet Math. Dokl.,* **1** 38–40 (1960)

[36] LIDSKIĬ, V.B. 'On the completeness of the system of eigen elements and adjoined elements of a compact operator' (Russian), *Dokl. Akad. Nauk SSSR,* **115** 234–236 (1957)

[37] LIDSKIĬ, V.B. 'Theorems on the completeness of the system of characteristic and adjoined elements of operators with discrete spectrum' (Russian), *Dokl. Akad. Nauk SSSR,* **119** 1088–1091 (1958)

[38] LIDSKIĬ, V.B. 'Conditions for the completeness of the system of root subspaces for non-self-adjoint operators with discrete spectrum' (Russian), *Trudȳ Moskva. mat. Obshch.* **8** 83–120 (1959)

[39] LIDSKIĬ, V.B. 'Non-self-adjoint operators with a trace' (Russian), *Dokl. Akad. Nauk SSSR.,* **125** 485–487 (1959)

[40] LIDSKII, V.B. 'Summation of series over the principal vectors of non-self-adjoint operators' (Russian), *Dokl. Akad. Nauk SSSR,* **132** 275–278 (1960)
English translation: *Soviet Math. Dokl.,* **1** 540–543 (1960)

[41] LIDSKIĬ, V.B. 'Summability of series in terms of the principal vectors of non-self-adjoint operators' (Russian), *Trudȳ Moskva. mat. Obshch,* **11** 3–35 (1962)

[42] LIVŠIC, M.S. 'On the spectral decomposition of linear non-self-adjoint operators' (Russian), *Mat. Sb.,* **34** (76) 145–199 (1954)
English translation: *Am. math. Soc. Transl. (2),* **5** 67–114 (1957)

[43] LOVITT, W.V. *Linear integral equations,* (Dover, New York, 1950)

[44] MACAEV, V.I. 'On a class of completely continuous operators' (Russian), *Dokl. Akad. Nauk SSSR,* **139** 548–551 (1961)
English translation: *Soviet Math. Dokl.,* 2^2 972–975 (1961)

[45] MACAEV, V.I. 'Volterra operators produced by perturbation of self-adjoint operators' (Russian), *Dokl. Akad. Nauk SSSR,* **139** 810–813 (1961)
English translation: *Soviet Math. Dokl.,* 2^2 1013–1016 (1961)

[46] MACAEV, V.I. 'A method of estimation of resolvents of non-self-adjoint operators' (Russian), *Dokl. Akad. Nauk SSSR,* **154** 1034–1037 (1964)
English translation: *Soviet Math. Dokl.,* **5** 236–240 (1964)

[47] MACAEV, V.I. 'Several theorems on completeness of root subspaces of completely continuous operators' (Russian), *Dokl. Akad. Nauk SSSR,* **155** 273–276 (1964)
English translation: *Soviet Math. Dokl.,* **5** 396–399 (1964)

[48] MERCER, J. 'Functions of positive and negative type, and their connection with the theory of integral equations', *Phil. Trans. R. Soc.* Ser. A., **209** 415–446 (1909)

[49] VON NEUMANN, J. and SCHATTEN, R. 'The cross-space of linear transformations'. III, *Ann. Math.,* **49** 557–582 (1948)

[50] PLEMELJ, J. 'Zur Theorie der Fredholmschen Funktionalgleichungen', *M. Math. Phys.* **15** 93–128 (1904)

[51] RINGROSE, J.R. 'Super-diagonal forms for compact linear operators', *Proc. Lond. math. Soc.* (3), **12** 367–384 (1962)

[52] RINGROSE, J.R. 'On the triangular representation of integral operators', *Proc. Lond. math. Soc.* (3), **12** 385–399 (1962)

[53] RUSTON, A.F. 'On the Fredholm theory of integral equations for operators belonging to the trace class of a general Banach space', *Proc. Lond. math. Soc.* (2), **53** 109–124 (1951)

[54] RUSTON, A.F. 'Direct products of Banach spaces and linear functional equations', *Proc. Lond. math. Soc.* (3), **1** 327–384 (1951)

[55] SAHNOVIC, L.A. 'The reduction of non-self-adjoint operators to triangular form' (Russian), *Izv. vȳssh. ucheb Zaved. Matematika* No. 1 (8), 180–186 (1959)

[56] SAHNOVIC, L.A. 'A study of the 'triangular form' of a non-self-adjoint operator'. (Russian), *Izc. vyssh. ucheb Zaved Matematika* No. 4 (11), 141–149 (1959)

[57] SMITHIES, F. 'The Fredholm theory of integral equations', *Duke math. J.*, **8** 107–130 (1941)

[58] SMITHIES, F. *Integral equations* (Cambridge, 1958)

[59] TAYLOR, A.E. *Introduction to functional analysis* (Wiley, New York, 1958)

[60] TITCHMARSH, E.C. *The theory of functions* (Oxford, 1950)

Index of Notation